VOLUME 11 ISSUE 1 2012

International Review of **Economics Education**

International Review of Economics Education

An international journal that promotes research into effective learning and teaching practice in economics higher education

The International Review of Economics Education is dedicated to enhancing learning and teaching in the higher education economics community. It provides a forum for high quality research in the areas of curriculum design, pedagogy, assessment, innovation and evaluation. The journal seeks to promote critical dialogue on educational theory and practice in economics and to demonstrate the relevance of research to good professional practice.

Design:
Kelvin Reeves Design (0117 914 7517)

With Thanks to the Higher Education Academy for providing financial support for this edition.

Submissions
Submissions should be sent in electronic form and should be uploaded through our Web interface. Further details of the submission procedure are available at:
www.economicsnetwork.ac.uk/iree/submit.htm

Ashley Lait, Managing and Production Editor
IREE, Economics Network
University of Bristol
8-10 Berkeley Square
Bristol, BS8 1HH, UK

Telephone: +44 (0) 117 331 4333
Email: ashley.lait@bristol.ac.uk

Contents

Editorial Issue 11.1

Peter Davies, Ross Guest and David McCausland

The papers in this issue address the impact of incentives and teaching on students' understanding. To what extent should economics educators expect incentives and/or teaching methods to improve students' understanding? Extrinsic or intrinsic incentives may affect students' effort and also the way they approach learning. Teaching methods may affect the fruitfulness of students' efforts.

Chen and Lin investigate students' use of online lectures as supplements – and to some degree – as substitutes for attending classes. Given evidence of the association between class attendance and grades, the relationships between class attendance, viewing online lectures and grades are of considerable interest. The associations found in their data are consistent with a positive effect of use of online lectures. As Chen and Lin conclude, this provides grounds for further studies which examine whether there is a causal effect.

More than ever, it is important for us as teachers to motivate and inspire students to want to learn economics. This is true more than ever because students can access so many alternative sources of economics content and learning resources available on the web, instead of attending our classes. If motivating students is so critical nowadays, we should be careful of potential de-motivators such as an overemphasis on grades rather than the intrinsic value of learning, which can occur according to the educational psychology literature. In this issue, Hadsell and MacDermott explore attitudes of economics faculty in the U.S. towards the importance of grades. They find that faculty like to emphasise grades in an attempt to motivate students. The likelihood that this is counterproductive for a good number of students is a concern. The authors suggest several strategies for de-emphasising grades in teaching.

Kneppers and colleagues investigate the way in which teaching helps students to build an integrated understanding of the subject. They distinguish between instances when students make connections between different abstract economic ideas and instances when students make a connection between an abstract idea and a particular context. Their results, from a comparison of two teaching approaches, emphasise the complexity of these learning processes. Teaching which encourages students to make connections makes a lot of sense, and it is difficult to see how teaching which treats the subject as a set of isolated ideas will be of much help to students. Still, as this research shows, we still have much more to learn about how students build a coherent understanding and how this coherence is affected by teaching.

Students' peers also have the potential to motivate and demotivate. Teasing out this effect empirically is, however, difficult. Contreras, Badua and Mitchell do this by apply a two stage least squares methodology. They find that "high ability" students have a positive effect on their "high ability" peers but a negative effect on "low ability" peers – the latter they attribute to an "intimidation factor". This finding adds grist to the mill of the age-old debate in education about streaming of classes by academic ability.

The use of classroom experiments, simulations and games - particularly online applications - is perhaps the fastest growing field in economics education. This is reflected in several recent articles in IREE including our special issue (9.2) in 2010. In the current issue Kennedy presents an online simulation to illustrate the concept of moral hazard arising from health insurance. In the game students typically exhibit moral hazard in that they seek health care more frequently and at a higher level when insured. Kennedy invites readers to contact him to obtain the software and assistance with importing the files.

Do Supplemental Online Recorded Lectures Help Students Learn Microeconomics?*

Jennjou Chen and Tsui-Fang Lin

Abstract

With the increasing popularity of information technology in higher education, it has become important to study how students use new technologies and how effective these methods are. This study sheds light on the relationship between the use of online recorded lectures and exam performance of students in the case of microeconomics. The study uses a rich panel data set covering Taiwanese students. Our results show that those who skip more classes and males are more likely to use online recorded lectures. As may be expected, most students access online recorded lectures just before exams, rather than immediately after lectures. Our fixed effects model shows a significant and positive relationship between students' use of online recorded lectures and their grades. On average, performance improvement attributable to the use of online supplements is close to 4 percentage points. In addition, watching online recorded lectures just before an exam increases students' performance by 3 to 5 percentage points.

JEL classification: A22, E32, E52

*We are grateful to three anonymous referees and seminar participants at the 2010 ASSA meetings for their thoughtful comments and suggestions. Financial support provided by Taiwan's National Science Council through grant NSC 98-2410-H-004-055-MY2 is gratefully acknowledged. All errors are our own.

1. Introduction

Many different technologies are incorporated into the teaching of microeconomics courses, including PowerPoint, chalk-and-talk, online discussion boards. However, it is not uncommon for students to face difficulties the first time they are introduced to complicated mathematical equations and graphs in the classroom. In the past, students sought help from their peers, teaching assistants and instructors, during or after the lectures. With advanced information technology now available, students are able to obtain extra help through e-learning. Nowadays, lectures can be recorded effortlessly through screen recording software while the instructors teach the subject. After the class, students have access to the recorded lectures for review and for exam preparation, any time and any where. In addition, the online recorded videos allow students to view the explanation of class materials as many times as needed. With the increasing popularity of information technology in higher education, it has become important to study how students use new technologies and how effective these methods are.

In spite of the importance of assessing the effectiveness of information technology on students' learning outcomes, only a few studies have explored such subjects specifically in the discipline of economics. For example, Agarwal and Day (1998) studied the effects of implementing internet enhancements, such as email, discussion lists and web information access, on students' learning. They found that the group that

had enhancements available performed better in exams. However, a study by Brown and Liedholm (2002) revealed the danger in relying too heavily on strictly online sources. They compared students' grades in traditional and online courses, and found that those who studied only online secured lower grades in exams. These two contradicting studies highlight the importance of studying how students are using the online materials as well as the result of such usage.

Two more recent studies conducted by Flores and Savage (2007) and Savage (2009) explored students' demand for supplemental online recorded lectures and the effects of such information technology enhancement on students' exam performance. Flores and Savage (2007) estimated students' willingness to pay for streaming lecture video and showed a positive correlation between video use and exam grades. Savage (2009) found that students exposed to downloadable lecture videos perform better on the exam but the positive effect is not statistically significant. However, both studies lack detailed information about students' use of supplemental online recorded lectures; their main conclusions are drawn from comparisons between the treatment and the control groups.

The main objective of this paper is to investigate the relationship between the use of a recent internet enhancement, online recorded lectures, and students' performance in exams. We have used a unique panel data set with details of students' usage of online resources. There are several channels that may affect students' learning. For example, in the case where students use internet enhancement as a substitute for live classroom learning, they benefit from streaming lecture videos but they lose the benefits of classroom interaction. If the provision of supplemental online lecture video does not significantly lower students' attendance rates then this method could complement their learning.

In this paper, we aim to answer the following three questions. First, if and when, do students adopt the e-learning approach and watch online recorded lectures? Second, what is the relationship between students' characteristics and their use of online recorded lectures? Third, does watching online recorded lectures improve students' exam performances? The main contribution of this study is to provide useful insights into the use of information technology in learning microeconomics and to explore the relationship between e-learning styles and students' performance in exams.

When assessing the impact of watching recorded lectures on students' exam performances, potential endogeneity problems arise, causing the ordinary least squares (OLS) estimates to be biased. For instance, in a case where students with strong motivations are more likely to use supplemental online resources and have better grades, the OLS estimates could overestimate the effectiveness of online recorded lectures. To address part of the potential endogeneity problem, we employ a fixed effects model to control for individual time-invariant heterogeneity and exam difficulty. As a result, we are able to obtain better estimates.

In the fixed effects model, variations of usage of online recorded lectures help us to identify the online lecture effects more accurately relative to the OLS estimates. Our fixed effects estimation results show that, on average, accessing supplemental online recorded lectures improves students' grades by 4 percentage points. In addition, the time when online recorded lectures are used matters in students' exam performances. On average, watching online recorded lectures soon after the live lectures does not improve students' grades, while watching them within the week before exams does enhance students' grades.

The details of the data set will be discussed in the next section. In the third section, the statistical models will be presented. The estimation results are then reported and the conclusion is provided in the final section.

2. Data

We conducted a survey of 312 students who took the intermediate microeconomics course at an elite public university in Taiwan in the fall semester of 2007 and spring semester of 2008. In our sample, three microeconomics courses were taught by the same instructor, and each contained 12 three-hour class meetings during the study period.

PowerPoint presentations were used in all class meetings. Each lecture was then recorded, using PowerCam software,[1] while the instructor was teaching the classes. Recorded lectures were uploaded to course websites immediately after each class meeting. The course website was located at the e-learning server provided and maintained by the university. Students needed to use their university email IDs and passwords to log on to the server and access course related information. The server then recorded students' use of online course materials.

It is important to note that students could only watch recorded lecture videos on the web and could not download the files to their personal computers. This is different from most available online recorded lectures that can be downloaded to students' personal computers. As a result, we were able to collect information on when students logged on to the server, whether or not they checked a specific course website, and most importantly, which videos were watched by the students and when. All 312 students' performances for each question in the exams were linked to their use of the online recorded lectures. For instance, we were able to observe whether or not, and how many times, a student accessed recorded files and how well each student did on the corresponding exam questions. In addition, information on students' attendance records corresponding to particular exam questions was also collected.

Exam questions included multiple choice questions and essay questions. When the instructor devised the exams, each question consisted of three parts: the question itself, an answer that was key to the question, and corresponding specific lectures and chapters in the textbooks during the sample semester. Thus, we were able to link students' performance in each question to their use of online recorded files.

In this paper, the dependent variable is students' grades, which are measured by the percentage of correctness in each answer. The percentage of correctness of an answer is defined as credits awarded, divided by the maximum credits for that particular question.

A set of variables were constructed to represent usage of online recorded lectures. In order to be sure that the respondent had accessed recorded lectures for review and exam preparation purposes, all variables refer to files accessed for at least five minutes. Definitions of these variables are as follows:

1. *Times*: the number of times a student accessed recorded lectures corresponding to a particular exam question.

2. *Watch Video*: whether or not the respondent had accessed online recorded files corresponding to a particular exam question.

3. *Total Minutes*: total time, in minutes, the respondent spent watching recorded lectures.

4. *Before the Exam*: whether or not the respondent had accessed recorded lectures less than a week before the exam.

[1] PowerCam is software developed by FormosaSoft Corporation; details of PowerCam software can be found at: http://blog.powercam.cc/en/.

5. *After the Lecture*: whether or not the respondent had accessed recorded lectures within a week after a lecture.

Our data do not allow us to empirically test whether the availability of online recorded lectures discouraged students from attending class lectures given that they could catch missed lectures later by accessing recorded lectures on the web. The main reason is that we cannot observe this group of sample students' attendance behaviour before the internet enhancement became available. In order to take into account the effect of attendance on exam performance, we also included students' attendance records corresponding to each exam question.

Table 1 presents the average attendance and average semester grades by students' use of online recorded lectures. The average attendance rate is 77.86%. According to the instructor's records, 77.86% is not significantly different from those in past semesters when the online recorded lectures were not available. Thus, it seems that the availability of online recorded lectures does not change students' attendance behaviour.

One-third of students in our sample had accessed online recorded lectures in the past. We found that there is no linear relationship between watching supplemental online video files and attending classroom lectures. For the students who never watch the videos, the average attendance rate is 77.93%, which is close to the overall average. The students who access online recorded lectures 6-10 times have the highest attendance rate, 85.91%. The average attendance rate for students with the highest frequency of viewing online recorded lectures is 51.43%. In terms of semester grades, students who never use online recorded lectures have the highest average semester grades at 72.25 points out of a total 100 points.

These results seem to suggest that the provision of supplemental online recorded lectures does not help students learn microeconomics. A possible reason could be that less motivated students might not attend lectures often when online recorded lectures are available to them. They might use these videos as substitutes for formal classroom lectures. Since they are less motivated, they might not care as much about what grades they get. The other potential explanation could be that students are not taught how to use the online recorded lectures effectively since the technology is new to them. Without taking into account other covariates and students' unobserved heterogeneity, we observe a negative association between online lecture usage and students' grades from Table 1. Below, a statistical model is presented to better estimate the relationship between students' use of online recorded lectures and their exam performances.

Table 1: Average class attendance and semester grades (by the number of times a student watches an online recorded lecture)

Number of Times	Number of Students	Average Class Attendance	Average Semester Grades
0	211	0.7793	72.25
1–5	33	0.8121	68.14
6–10	22	0.8591	62.33
11–15	25	0.7120	70.97
16–20	14	0.8143	69.16
21+	7	0.5143	62.15
Total	312		
Mean		0.7786	70.65

Note: The average attendance rates are computed as described below. First, we computed each student's average attendance rate for the semester (i.e. lectures attended/total lectures). Then we computed the average attendance rate for each group.

3. Statistical models

This study uses a micro level data set to explore the relationship between students' use of online recorded lectures and their performance in exams. As described earlier, the OLS model could suffer from potential endogeneity bias. A fixed effects model is, therefore, employed to address part of the endogeneity problem.

A linear model describing the relationship between a student's exam performance and various learning inputs variables is shown as follows:

$$y_{ij} = \eta r_{ij} + \lambda t_{ij} + \alpha_i + \gamma_j + \varepsilon_{ij}, \quad i = 1, 2, \ldots, I, j = 1, 2, 3, \ldots, J \qquad (1)$$

I denotes the total number of students and J denotes the total number of exam questions. y_{ij} corresponds to student i's observed exam performance on question j. r_{ij} refers to three *Recorded Lecture* variables: *Watch video*, *Watch video right after lecture*, *Watch video right before exam*. The definitions of these variables are discussed in the previous data section. η captures the association between use of recorded lectures and grades, the major interest of this paper. t_{ij} represents student i's attendance record for question j when the materials of question j were taught. α_i represents student i's time-invariant individual effect, γ_j represents question j's specific effect, and ε_{ij} is a random disturbance term.

We included students' attendance status in equation (1) since it is important in determining students' exam performance. We added the attendance variable and the interaction terms between attendance and viewing behaviour as control variables in our model to better estimate the relationship between online lecture usage and exam performance.

Table 2 presents summary statistics of variables. Since 312 students' exam performances were observed for each exam question, there are 13,490 observations. The first panel shows summary statistics for the full sample. The second panel shows summary statistics when attendance is coded as zero. And the third panel shows summary statistics when attendance is coded as one.

Table 2: Summary statistics

Variables	N	Mean	Standard Deviation	Minimum	Maximum
		Full Sample			
Attendance	13490	0.7599	0.4272	0.0000	1.0000
Grades	13490	0.6736	0.4212	0.0000	1.0000
Male	13490	0.4098	0.4918	0.0000	1.0000
Times	13490	0.3202	0.7893	0.0000	7.0000
Watch Video	13490	0.1834	0.3870	0.0000	1.0000
Total Minutes	13490	26.722	79.790	0.0000	1053.0
Before the Exam	13490	0.1645	0.3707	0.0000	1.0000
After the Lecture	13490	0.0634	0.2437	0.0000	1.0000
		Attendance = 0			
Attendance	3239	0.0000	0.0000	0.0000	0.0000
Grades	3239	0.6300	0.4390	0.0000	1.0000
Male	3239	0.5536	0.4972	0.0000	1.0000
Times	3239	0.4594	0.9510	0.0000	7.0000
Watch Video	3239	0.2485	0.4322	0.0000	1.0000
Total Minutes	3239	40.577	104.40	0.0000	834.00
Before the Exam	3239	0.2186	0.4134	0.0000	1.0000
After the Lecture	3239	0.1099	0.3128	0.0000	1.0000
		Attendance = 1			
Attendance	10251	1.0000	0.0000	1.0000	1.0000
Grades	10251	0.6874	0.4145	0.0000	1.0000
Male	10251	0.3644	0.4813	0.0000	1.0000
Times	10251	0.2762	0.7254	0.0000	7.0000
Watch Video	10251	0.1628	0.3692	0.0000	1.0000
Total Minutes	10251	22.345	69.680	0.0000	1053.0
Before the Exam	10251	0.1474	0.3545	0.0000	1.0000
After the Lecture	10251	0.0487	0.2152	0.0000	1.0000

Note: Students are repeatedly observed by the questions they answered in exams. Therefore, the number of observations is 13,490 rather than 312 in Table 1.

By comparing the summary statistics in the second and third panels, we found that attendance is positively associated with students' performances. Male students, relative to female students, are more likely to skip lectures. Most importantly, we found that students are more likely to access recorded lectures if they skip lectures.

Table 3 shows correlations among key variables. We found that grades are positively correlated with attendance. The set of *Recorded Lecture* variables is negatively associated with both attendance and

grades without controlling for other factors. In addition, male students, relative to their female counterparts are more likely to access online recorded lectures.

Table 3: Pearson correlation coefficients for attendance, grades, online video usage, and other variables

Variables	Attendance	Grades	Male	Times	Watch Video	Total Minutes	Before the Exam	After the Lecture
Attendance	1							
Grades	0.0582**	1						
Male	–0.1643**	–0.0423**	1					
Times	–0.0991**	–0.0007	0.0164**	1				
Watch Video	–0.0946**	–0.0142*	0.0165**	0.8559**	1			
Total Minutes	–0.0976**	–0.0121	–0.0139	0.7531**	0.7059**	1		
Before the Exam	–0.0820**	–0.0198**	0.0319**	0.7892**	0.9362**	0.6508**	1	
After the Lecture	–0.1073**	–0.0666**	–0.0051	0.4171**	0.5489**	0.3599**	0.4935**	1

Note: Number of observations is 13,490. "**" at the 5% and "*" at the 10% significance level.

4. Estimation results

Estimation results of both OLS models and fixed effects models are presented in Table 4. Clustered standard errors are computed, using the textbook chapter as the cluster. The following main conclusions are based on the fixed effects model since we relied on variations within each of the students' viewings to identify the online recorded lectures' effects.

Table 4: The effects of supplemental online recorded lectures

Independent Variable	OLS							Fixed Effects						
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
(1) Attendance	0.0591**	0.0616**	0.0611**	0.0672**	0.0655**	0.0706**	0.0709**	0.0151	0.0170	0.0159	0.0189	0.0192*	0.0212*	0.0217*
	(0.0127)	(0.0127)	(0.0127)	(0.0154)	(0.0138)	(0.0157)	(0.0148)	(0.0114)	(0.0110)	(0.0109)	(0.0121)	(0.0109)	(0.0118)	(0.0119)
(2) Watch video		0.0235**							0.0389**					
		(0.0080)							(0.0111)					
(3) Watch video right after lecture			0.0319	0.0351*	0.0631**	0.0272	0.0392**			0.0007	0.0019	0.0218	-0.0017	0.0126
			(0.0192)	(0.0195)	(0.0171)	(0.0201)	(0.0179)			(0.0204)	(0.0203)	(0.0179)	(0.0198)	(0.0170)
(4) Watch video right before exam			0.0042	0.0229	0.0041	0.0407*	0.0353			0.0340*	0.0421**	0.0336*	0.0531**	0.0468**
			(0.0145)	(0.0189)	(0.0145)	(0.0226)	(0.0225)			(0.0187)	(0.0201)	(0.0188)	(0.0194)	(0.0210)
(5) Interaction Term (1)*(2)				-0.0289							-0.0130			
				(0.0172)							(0.0167)			
(6) Interaction Term (1)*(3)					-0.0495**		-0.0181					-0.0338**		-0.0218
					(0.0187)		(0.0145)					(0.0156)		(0.0216)
(7) Interaction Term (1)*(4)						-0.0484**	-0.0414**						-0.0254	-0.0173
						(0.0180)	(0.0175)						(0.0187)	(0.0223)
Hausman Test (OLS vs. Fixed Effects)								35.320	44.610	9.0900	10.760	8.1900	8.6900	8.3900
R-squared	0.2763	0.2766	0.2766	0.2768	0.2768	0.2770	0.2770	0.3596	0.3602	0.3601	0.3601	0.3602	0.3602	0.3602
Sample Size	13,490	13,490	13,490	13,490	13,490	13,490	13,490	13,490	13,490	13,490	13,490	13,490	13,490	13,490

Note: The dependant variable is a percentage of the correctness of an answer. Exam question dummies are included in all regressions. The row of Hausman test reports the statistics for OLS against fixed effects models. "**" is significant at 5% and "*" is significant at 10% type I error level. White (1980) robust standard errors are in parentheses; standard errors also robust to clustering in square brackets.

Attendance has a significant and positive effect on students' learning outcomes in most models. In the OLS model, the grades improvement attributable to attendance ranges from 5.91 percentage points to 7.09 percentage points; in the fixed effects model, the grades improvement attributable to attendance ranges from 1.51 percentage points to 2.17 percentage points. This finding is in line with those in prior literature using either secondary data (Marburger, 2001, 2006; Stanca, 2006) or experimental data (Chen and Lin, 2008).

Another one of our key independent variables, whether or not students utilise online recorded lectures (*Watch Video*), also has a positive correlation with students' performance in exams. The association is statistically significant in both OLS and fixed effects models. It is important to note that after controlling for time-invariant individual effect and question specific effect, the fixed effects model shows that watching online recorded lectures has a greater impact. The improvement in grades associated with using online recorded lectures is close to 4 percentage points. This finding is consistent with results in some recent studies (Flores and Savage, 2007; Savage, 2009).

In addition, we found that when a student uses online recorded lectures matters in students' exam performances. Watching online recorded lectures right after the live lectures improves students' grades in some of the OLS models. However, the improvement becomes insignificant in the fixed effects model. In contrast, on average, watching recorded lectures in the week before an examination improves students' grades by 3 to 5 percentage points in the fixed effects model.

There are several plausible explanations for this finding. First, students may simply be cramming the material before exams. Second, students may become more serious about their studies because a test is approaching and use the online videos concurrently. Third, as students begin to study for tests, they may realise the gaps in their comprehension and rely on the videos to enhance their understanding.

If students are using online videos for cramming, then the availability of online recorded lectures helps students pass the exam and promotes the less desirable learning style which could hamper students' learning in the long term. Whereas if students are using the online recorded lectures concurrently with their notes, then the videos are complementing their studies. And if students are using the online recorded lectures to fill in the gaps in comprehension, then the online videos are acting as review sessions and help students' learning too.

Lastly, we also included several interaction terms to test whether the effect of accessing online recorded lectures differs with attendance records of students. We found that none of the interaction terms is statistically significant, after taking into account time-invariant heterogeneity and the question dummy.

5. Conclusion

With the increasing popularity of information technology in higher education, it is important to study how students use new technologies for learning different subjects, and the results of adopting information technology on their learning outcomes. This study, using a rich panel data set, sheds light on the relationship between the usage of online recorded lectures and exam performance, specifically in the case of Microeconomics and Taiwanese students.

Our results show that one-third of students in our sample had accessed online recorded lectures. In addition, males, and those students who skip relatively more classes, are more likely to use online recorded lectures. With regard to the timing of watching online recorded lectures, most students access them just before sitting exams, rather than after lectures.

After taking into account time-invariant heterogeneity and question-specific effects, we partially addressed the potential endogeneity bias resulting from the correlation between online lecture viewing

behaviour and unobserved heterogeneity. Our fixed effects model shows a significant and positive association between students' use of online recorded lectures and their exam performances. On average, the improvement associated with such online supplements is 4 percentage points, which is greater than the class attendance effect for the same group of students. The estimation results show that, after controlling for attendance status, the provision of supplemental online recorded lectures does help students learn microeconomics and improves their grades. For this group of students, those who use online resources relative to those who do not use online resources get better grades. This may imply that students could use online recorded lectures to review covered materials and that enhances their understanding of microeconomics.

Furthermore, the results also show that when students view the online lectures relative to when they take tests affects their exam performance. Accessing online recorded lectures just before an exam increases students' performance by 3 to 5 percentage points while watching online recorded lectures immediately after a class meeting does not show a significant effect in test scores. This might imply that the availability of online recorded lectures enhances students' exam performances and possibly increases the number of students who pass the exam by cramming the materials covered.

Many researchers are concerned about the correlation between lower attendance rates and the availability of internet enhancements since poor attendance rates usually mean students learn less too. This paper adds value to existing literature by investigating the usage of online recorded lectures and exploring the relationship between students' adoption of online technology and their academic performance. Future research needs to be done to better address the causal relationship between internet enhancement, attendance patterns and students' learning outcomes.

References

Agarwal, R. and Day, A. E. (1998). 'The impact of the internet on economic education, *Journal of Economic Education*', vol. 29(2), pp. 99–110.

Brown, W. B. and Liedholm, C. E. (2002). 'Can web courses replace the classroom in Principles of Microeconomics?', *American Economic Review*, vol. 92(2), pp. 444–48.

Chen, J. and Lin, T. F. (2008). 'Class attendance and exam performance: a randomized experiment, *Journal of Economic Education*, vol. 39(3), pp. 213–27.

Flores, N. and Savage, S. J. (2007). 'Student demand for streaming lecture video: empirical evidence from undergraduate economics classes', *International Review of Economics Education*, vol. 6(2), pp. 57–78.

Marburger, R. (2001). 'Absenteeism and undergraduate exam performance', *Journal of Economic Education*, vol. 32(2), pp. 99–110.

Marburger, R. (2006). 'Does mandatory attendance improve student performance?', *Journal of Economic Education*, vol. 37(2), pp. 148–55.

Savage, S. J. (2009). 'The effects of information technology on economic education,' *Journal of Economic Education*, vol. 40(4), pp. 337–53.

Stanca, L. (2006). 'The effect of attendance on academic performance: Panel data evidence for introductory microeconomics,' *Journal of Economic Education*, vol. 37(4), pp. 251–266.

White, H. (1980). 'A Heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity', *Econometrica*, vol. 48(4), pp. 817–38.

Contact details

Jennjou Chen
Department of Economics
National Chengchi University
Taipei City 116, Taiwan
Email: jennjou@nccu.edu.tw

Tsui-Fang Lin
Department of Public Finance
National Taipei University
New Taipei City 237, Taiwan
Email: tflin@gm.ntpu.edu.tw

Faculty Perceptions of Grades: Results from a National Survey of Economics Faculty*

Lester Hadsell and Raymond MacDermott

Abstract

Results from a survey of U.S. economics faculty (816 responses) indicate the extent to which grades are emphasised in their classes. We measure learning- and grade-orientations and relate our findings to empirical research in economics and educational psychology. We find agreement among economics faculty on a broad range of grade-oriented attitudes and behaviours. We note differences between views of economics faculty and empirical research on several key topics. Free-form comments indicate a concern with grade distributions, the influence of grades on student evaluations of teaching, and grade inflation.

JEL classification: A2

* We thank David Ring and Jaime McGovern for comments on earlier versions of the survey; seminar participants at Washington and Lee University, SUNY Oneonta and UNC-Wilmington, and conference participants at the ASSA meetings in San Francisco (2009) for comments on preliminary results; Diana Moseman for her assistance with administration of the mailing and survey; and several student assistants at Oneonta and VMI for assistance in collecting email addresses. We also thank the hundreds of colleagues who completed the survey. We gratefully acknowledge funding for data collection from SUNY College at Oneonta and VMI.

1. Introduction

Relatively little is known about the views of economics faculty concerning grades despite the important role grades play in educational settings. Studies by Becker and Watts (1996, 2001a, 2001b) and Schaur *et al.* (2008) provide a good idea of which assessment tools economics faculty use. We also have information about best practice in assessment (Walstad, 2006). But little is known of faculty attitudes or behaviours toward grading or their perceptions of how students react to grades. This lack of knowledge is particularly startling given the apparent importance placed by economics faculty on grades (as evidenced by responses to our survey) and the influence that grading policies have on student learning outcomes and effect.

We know from research in educational psychology that the grading policies of teachers affect students' performance (Meece *et al.*, 2006). Policies that de-emphasise grades and promote mastery (learning) goals generally are positively associated with desired student outcomes, including increased learning (e.g. Harter, 1978; Moeller and Reschke, 1993), effort (Grolnick and Ryan, 1987; Ames and Ames, 1991), help seeking (Karabenick, 2004; Linnenbrink, 2005), enjoyment (Pekrun *et al.*, 2006), and long-term interest (Butler, 1987; Harackiewicz *et al.*, 2000; Henderlong and Lepper, 2002; Senko and Harackiewicz, 2005).[1] In contrast, an emphasis on grades generally is associated with anxiety, hopelessness and shame (Linnenbrink, 2005; Pekrun *et al.*, 2006), effort withholding (Thompson, 1994; Urdan *et al.*, 1998;

[1] Subject interest was reported by economics alumni as the second most important determinant in choosing to major in economics – behind positive experience in the principles class (Allgood *et al*., 2004).

Thompson and Perry, 2005), and preference for less challenging tasks (Harter, 1978). Of course, some students respond favourably to conditions that emphasise grades (Deci *et al.*, 1999; Cherry and Ellis, 2005; Betts and Grogger, 2003) but, as we discuss below, even these studies report students who are left behind.

This paper reports views on grades and grading policies of economics faculty across the U.S., adding to our knowledge of the extent to which economics faculty emphasise grades. Responses to our survey indicate that there is consensus among economics faculty on a range of grade-oriented views and behaviours and that economics faculty generally place great value on grades. We find that economics faculty tend to view grades as good motivators and are concerned about grading standards. Further, 40 per cent of faculty indicate they are influenced either 'some' or 'a great deal' by departmental or college expectations when they grade students' work. Free-form comments indicate concerns with grade inflation and use of student evaluations (which are seen to depend in part on grades) in promotion, tenure and pay decisions.

An examination of how economics faculty view and use grades is worthwhile given the extensive evidence from educational psychology. While the behavioural view of human motivation (e.g. Skinner, 1976) dominates economics, with its reliance on extrinsic motivation, cognitive psychologists have demonstrated that intrinsic motivation is important as well. In this more expansive view, people have some level of natural curiosity, seek to resolve discrepancies between what they see and what they know, and have aspirations and varying degrees of need for achievement. A comprehensive view of academic motivation takes into account the intrinsic as well as extrinsic motivations of students and sets classroom policies accordingly. While extrinsic rewards such as grades are effective motivators for some tasks (tedious, repetitive), they are less effective – and can even be detrimental – when applied to other tasks (those that are inherently interesting). In the latter case, the tasks are said to have intrinsic value and the extrinsic rewards may crowd out that value, reducing the students' interest (recent work in labour economics explores crowd-out as well; e.g. Falk and Kosfeld, 2006; Frey, 1998). In short, an over-reliance on extrinsic rather than intrinsic motivation will often lead to suboptimal outcomes.[2]

Our survey of the views of grades of economics faculty is made in light of this extensive evidence. Our primary goal is to measure the degree to which economics faculty emphasise extrinsic motivation (grades) and intrinsic motivation ('learning' goals) in their classrooms so that economics faculty as a whole can critically evaluate classroom policies. In making this distinction between intrinsic and extrinsic incentives, we do not mean to imply that grades and learning are mutually exclusive. Nor do we want to suggest that grades do not supply useful information to students and others (potential employers, parents, admissions offices) on the level of achievement attained by students. It is reasonable to conclude that under a given grading structure, a higher grade reflects, on average, a higher level of knowledge or understanding.[3] We also do not argue that grades are necessarily competitive or ineffective incentives. Rather, we maintain that what the empirical findings from educational psychology tell us is that the level of learning can be influenced (shifted) by the grading environment perceived by the student, sometimes in unexpected ways, and as a consequence we need to be aware of what we as faculty think about grades and how we use them.

[2] Those wishing a comprehensive review may consult Ames and Ames (1991), Deci *et al*. (1999) and Meece *et al*. (2006).

[3] For an argument that grades are subject to measurement error see Kirschenbaum *et al.* (1971, pp. 54-6). For example, they discuss a study in which copies of two high-school (secondary) English exam papers were sent to English teachers at 200 high (secondary) schools who were asked to mark the papers as they normally would at their school. The scale was 100 points, with 75 as passing. For one paper scores ranged from 64 to 98, with an average of 88.2. For the other, the range was 50 to 97, with an average of 80.2. Based on written comments provided by the scorers, the differences between high and low scores appeared to be based on subjective evaluation of what makes a 'good paper'; on varying emphasis on neatness, spelling, and punctuation; and on general ability of the writer to communicate the paper's primary message. A second study was performed by the same researchers, with a geometry assignment instead. The range of scores increased to 67 (low: 28 per cent; high: 95 per cent). Based on my own limited observations of similar cases in which papers were judged by a group of faculty, even using a common grading rubric does not solve the problem. Furthermore, other studies have shown that the same papers graded by the same instructor, separated by time, can receive markedly different scores. Thus, grades generated by different faculty or at different times contain much noise and as such may be unreliable measures of knowledge.

In the next section, we describe the survey instrument and method. We then summarise and analyse the findings. A selection of free-form comments are offered in the following section. Our paper concludes with a few recommendations for incorporating the insights from this research into the teaching of economics.

2. The survey and method

The survey measures faculty views of learning and grades along a two-dimensional scale first developed by educational psychologists (Janzow and Eison, 1990; Eison *et al.*, 1993). Learning oriented (LO) attitudes and behaviours (five statements each) and grade oriented (GO) attitudes and behaviours (five statements each) comprise the 20 statement survey. Respondents use a five-point scale (1 = Strongly disagree/Never; 5 = Strongly agree/Always) to indicate level of agreement with the attitudinal statements (numbers 1 though 10) and frequency of use with the behavioural statements (numbers 11 through 20). The survey questions are shown in Table 2.

LO statements measure the extent to which faculty engage in attitudes or behaviours that have been identified in the educational psychology and education literatures as focusing students on learning (mastery). GO statements measure the extent to which faculty engage in attitudes or behaviours that focus students on grades (performance). Responses to the LO statements (2, 4, 7, 8, 10, 13, 14, 15, 16, 20) can be summed to form a total LO score. The remaining responses form the total GO score. Higher scores indicate greater LO or GO. Note that faculty can have (or lack) both LO and GO. Principal component analysis (on the data collected in this survey, and in prior surveys) supports the validity of this grouping of statements.

Each statement in the survey has empirical support in the educational psychology literature for its stated orientation. LO promotes collaboration, encourages improvement and provides choice. GO emphasises performance measurement, focus on the 'best and brightest' students, and competition. Faculty with higher LO scores tend to view grades as over-emphasised and overvalued and tend to be flexible in grading and the way they view disciplinary boundaries, while higher GO is associated with a concern with grade inflation and attention to the significance of grade point average (GPA) (Eison *et al.*, 1993).

The survey was administered online on the Oneonta College website and consisted of the 20 statements measuring learning and grade orientations plus an additional 11 demographic questions (contained in Appendix 1), which asked for information about the respondent and institution, and about such items as teaching loads, evaluation tools and weights of those tools when determining course grades at the undergraduate level. A request to complete the survey was emailed to 5915 members of 599 economics departments in the U.S.[4] Of the emails sent, 149 were returned as undeliverable. From the remaining 5766 recipients, 816 surveys were completed between 23 and 30 September 2008. The 816 responses represent a 14.2 per cent response rate. This response rate is similar to other recent national surveys of economics faculty (Schaur *et al.*, 2008).[5]

Table 1 (on the next page) provides a summary of the demographic and institutional data collected in our survey. These data indicate that 74 per cent of respondents were male, 40 per cent were full professors, and half were either associate (25 per cent) or assistant (24 per cent) professors, and the average number of years teaching was just more than 17.[6] Half of the respondents taught in PhD granting departments while 31 per cent taught in departments where the bachelors is the highest degree awarded. Average department size across the entire sample was 15.5 members.

[4] The email addresses were collected manually from individual department websites by the authors and several student assistants.

[5] The rate for Schaur *et al.* was 13.0 per cent in a similar mailing in 2005. They received 477 responses from a mailing to 3658 faculty using a private market mailing list – one from Market Data Retrieval (MDR). The advantage of MDR is that it identifies recipients by instructor specialisation, allowing researchers to determine response rates by specialisation. A disadvantage of MDR is its cost and one-time use policy (researchers do not actually have the email addresses – and cannot verify them).

[6] The percentage of females in the sample is generally reflective of the percentage in the profession (AEA CSWEP, 2008).

Table 1: Summary statistics (counts, unless otherwise noted)

Gender	Male	Female	N/R				
	581	191	8				
	74%	24%	1%				
Academic Rank	Full	Associate	Assistant	Other			
	316	197	185	83			
	40%	25%	24%	11%			
Years of teaching experience, average	17.51						
Number of faculty in department, average	15.48						
Highest degree offered	Ph.D.	MA	Bachelors	MBA	N/R		
	392	143	238	2	3		
	50%	18%	31%	0%	0%		
School in which department is housed	Agriculture	Business	Humanities	Liberal Arts	Public Policy	Science	N/R
	3	283	230	114	11	118	8
	0%	37%	30%	15%	1%	15%	1%
Influence of departmental grading expectations (Survey item #31)	A great deal	Some	Very little	Not at all			
	78	255	210	224			
	10%	33%	27%	29%			
Sections taught, number of students (averages for those who taught each level)		Sections	Students				
	Principles	2.59	162.2				
	Intermediate	1.62	65.3				
	Upper	1.82	55.8				
	Masters	1.29	35.6				
	Ph.D.	0.91	12.1				

	Principles		Intermediate		Upper	
	Primary	Secondary	Primary	Secondary	Primary	Secondary
Multiple Choice Exam		9%	7%	5%	4%	3%
Short Answer Exam		6%		7%		8%
Homework	2%		3%		2%	
Papers	1%	7%	3%			
Quizzes	1%	9%	1%	3%	1%	2%
Class Participation	0%	3%	0%	3%	1%	

While the number of recipients and the method used for acquiring their addresses supports our belief that they are representative of the academic economists in the U.S. generally, we have no way of knowing for sure whether the same is true for the respondents.[7] Given the opportunistic nature of our sample, we cannot be certain that the respondents reflect the profession as a whole. One approach to evaluating the representativeness of a survey sample (i.e. test for non-response bias) is to compare the early and late responses (Bose, 2001; Oppenheim, 1966). The presumption is that late responses (rather than the early responses) are more similar to non-responses, so that any differences between early and late responses indicate a non-response bias. We examine the first 75 and last 75 responses in our sample, and perform a series of *t*-tests for differences in sample means (for each data series gathered). The results indicate no significant (p<0.05) difference between early and late responses, except for degree level of institution (MA were more likely to be late responders, PhD early) and number of graduate students and economics majors taught (early responders were more likely to teach more of both). Given that there were no significant differences in LO and GO scores, gender, years teaching, teaching in business schools, teaching principles, teaching intermediate, and teaching upper level, we conclude that non-response bias is not an issue in this study.

3. Survey results

Individual statements

Summary statistics of responses are shown in Table 2. These results show that faculty generally exhibited strong grade oriented attitudes (GOA). Their broad agreement to GOA statements generally indicates support for grades as incentives (statements 1, 5) and the validity of grades as a measure of performance or ability (3, 6, 9). Statements 1 and 5 are the most agreed upon statements in the survey: 91 per cent of respondents agreed that grades were useful tools for increasing student performance; 89 per cent agreed that regularly scheduled exams were necessary for students to be expected to learn. Economics faculty appear to have a strong inclination to believe in the effectiveness of grades as extrinsic motivators, despite the limitations of grades noted earlier.

Mixed reviews of the effectiveness of grades as an extrinsic motivator have also been reported in the recent economics education literature. Grove and Wasserman (2006), for example, find that freshmen score better on exams when assignments are graded but other students do not, and Betts and Grogger (2003) find that while tougher grading (in high school) is initially correlated with higher scores on standardised tests, the long-term effect on scores is negligible – with the exception of minority students, for whom the effects are negative. And in a study involving students in introductory microeconomics, Dickie (2006) finds that grade incentives appear to exert a negative influence that offsets the beneficial effect of classroom experiments.

[7] We used a single survey address to collect all the responses. In retrospect, we could have provided a separate survey for each Carnegie classification. That is, we could have created separate mailing lists by classification and directed recipients to a particular survey site depending on their classification. This would have provided response rates by type of institution. As it stands, we are not able to determine the response rate by classification. We do know that 50 per cent of the respondents indicated that the highest degree offered at their institution was the PhD whereas 54 per cent of all economics faculty (and presumably 54 per cent of the recipients) belong to PhD institutions (Kamath *et al.*, 2007). In our sample, 18 per cent belong to Masters institutions and 31 per cent to bachelors, whereas nationally the percentages are 30 and 10, respectively.

Table 2: Summary statistics of responses

#	Statement	Type	Ave	s.d.	(5)	(4)	(3)	(2)	(1)
	"Attitudes" statements 1-10				Strongly agree	Agree	Neutral	Disagree	Strongly disagree
1	Without regularly scheduled exams most students would not learn the material I present.	GOA	4.26	0.87	45%	44%	5%	5%	1%
3	I think college grades are good predictors of career success in later life.	GOA	3.43	0.76	3%	49%	36%	10%	1%
5	I think it is useful to use grades as incentives to increase student performance.	GOA	4.14	0.64	26%	65%	7%	2%	0%
6	I wish my colleagues across the campus were tougher graders.	GOA	3.96	0.84	30%	41%	25%	4%	0%
9	I worry about colleagues who are giving an ever increasing number of As and Bs.	GOA	3.93	0.92	29%	45%	18%	8%	1%
2	I think students should be encouraged to collaborate rather than compete.	LOA	3.54	1.01	16%	42%	27%	12%	4%
4	Students' concern about grades often interferes with learning in my classroom.	LOA	3.07	1.12	9%	32%	20%	33%	6%
7	I don't mind if students enroll in my classes under the pass/fail/audit options.	LOA	3.60	1.06	20%	40%	23%	13%	4%
8	I think my colleagues across campus place too much emphasis on using grades to motivate students.	LOA	2.43	0.79	1%	5%	40%	44%	10%
10	I would prefer teaching a course in which no grades were given rather than a typical graded course.	LOA	2.26	1.09	5%	10%	20%	40%	27%
	"Behaviors" statements 11-20				Always	Often	Sometimes	Seldom	Never
11	I set grading standards that are designed primarily to challenge the brightest students in my classes.	GOB	3.08	1.07	8%	28%	37%	18%	9%
12	I emphasize in my conversations with students the importance of studying to obtain 'good grades.'	GOB	2.71	1.18	7%	21%	26%	29%	18%
17	I orient my teaching style (e.g., content, pace, difficulty level) to satisfy the needs of upper level students and hope that the others can keep up.	GOB	2.74	0.95	4%	16%	38%	34%	8%
18	I encourage students to focus primarily on their studies and to limit their participation in extracurricular activities which might jeopardize their GPA.	GOB	1.78	0.96	1%	4%	15%	29%	50%
19	I tell students that competition for grades prepares them for the competitive nature of adult life.	GOB	1.60	0.91	1%	4%	11%	21%	63%
13	I allow students the opportunity to choose among alternative assignments as a way to enhance motivation.	LOB	2.12	1.05	2%	9%	24%	31%	35%
14	I encourage students to raise questions in class that are topic-related but which also go beyond the scope of the tests which I prepare.	LOB	4.26	0.77	44%	41%	14%	2%	0%
15	I am willing to make exceptions to stated grading criteria when unusual circumstances arise.	LOB	2.83	1.05	8%	15%	36%	32%	8%
16	I design course assignments that encourage students to read outside my discipline.	LOB	2.65	1.03	5%	15%	33%	35%	12%
20	I reward student improvement and growth by weighing the students' progress in my grading system.	LOB	2.81	1.18	8%	20%	32%	22%	17%

GOA = Grade Oriented Attitude statement
LOA = Learning Oriented Attitude statement
GOB = Grade Oriented Behavior statement
LOB = Learning Oriented Behavior statement

The responses to statement 6 (and 9) indicate that economics faculty think faculty in other disciplines are easy graders. Grade inflation is a concern as well. Further, 89 per cent agreed with statement 3, 'I think college grades are good predictors of [career] success in later life', even though research findings indicate a tenuous relationship between grades and future career success (Cohen, 1984; Baird, 1985; Davidson and Lewis, 1997).[8] Thus, economics faculty place a heavy emphasis on grading and grades.

The next section of Table 2 shows that economics faculty exhibit mixed learning oriented attitudes (LOA), with strong LOA in their responses to statements 2 and 7, which generally de-emphasises the role of grades, but anti-LO in their responses to statements 8 and 10, which indicate a belief that grades are a necessary motivator for students (consistent with the strong GOA). A majority favour collaboration over competition (statement 2) and enrollment under the pass/fail/audit option (statement 7), both of which have been shown to promote positive learning outcomes. On the other hand, a majority disagree that faculty in other disciplines place too much emphasis on using grades to motivate students (mirroring the responses for statements 6 and 9).

In response to statement 4 ('Students' concern about grades often interferes with learning in my classroom'), a large minority of faculty identified concern with students' focus on grades (41 per cent agreed or strongly agreed to the statement). As 41 per cent agreed, we can reasonably conclude that for many, but not most, faculty students' concern about grades is a problem.[9]

Given the strong GOA noted above, it is surprising that respondents exhibit very little overt grade oriented behaviour (GOB), as shown in the next set of responses in Table 2. Teaching style and grading standards are not designed to accommodate more capable students at the expense of the less capable (statements 11 and 17). Further, faculty do not emphasise the importance of grades in conversations with students (statements 12, 18 and 19).

Respondents also score low on the LO behaviour (LOB) scale. Faculty allow students little choice in completing assignments, at least in an effort to enhance motivation (statement 13). As discussed in a previous section, choice has been shown to be an important factor in student motivation. Responses to statement 13 indicate that economics faculty provide little opportunity to 'choose among alternative assignments as a way to enhance motivation'. Anderman and Midgley (1998) note that allowing some degree of control over learning by giving students choices between different assignments does not mean teachers must relinquish control of the classroom: 'Even small opportunities for choice, such as whether to work with a partner or independently,' give students a greater sense of autonomy.

Two out of every five faculty indicated that improved performance is not weighed in grading decisions (statement 20). Further, many faculty are not willing to make exceptions to stated grading criteria when unusual circumstances arise (statement 15), perhaps concerned with the implications of breaking a 'contract' – as the syllabus is often viewed as representing. Also faculty are reluctant to encourage students to read from outside of the economics discipline (statement 16), despite the inherent multidisciplinary nature of economics. In many respects this is understandable, given time constraints and training typically provided to economics faculty.

[8] Davidson and Lewis (1997) find that less prepared medical school applicants earned lower GPAs in medical school but did equally as well in their careers as those with better GPAs or scores going into medical school. Cohen (1984) conducts a meta-analysis of 108 studies correlating grade average in college to various criteria of adult achievement or success. His conclusions 'may be somewhat discouraging to those placing great importance on grades and their predictive power. It seems that how well a student does in college relates only marginally with success in a career' (p. 292). Professors, who typically have earned high GPAs during their many years of academic study, may be expected to believe that grades are strong indicators of future success. Nonetheless, the evidence is much less certain.

[9] The 39 per cent who disagreed or strongly disagreed do not cancel out the 'agrees'. If two out of five faculty indicate it is a problem then the fact that it's not a problem for the others is not central.

Overall, economics faculty are mostly in agreement (as measured by standard deviation of the responses) to GO statements more so than LO statements (8 of the 10 most agreed upon statements are GO). In sum, the survey responses suggest that economics faculty place heavy emphasis on the extrinsic motivating qualities of grades.

Demographic and institutional influences

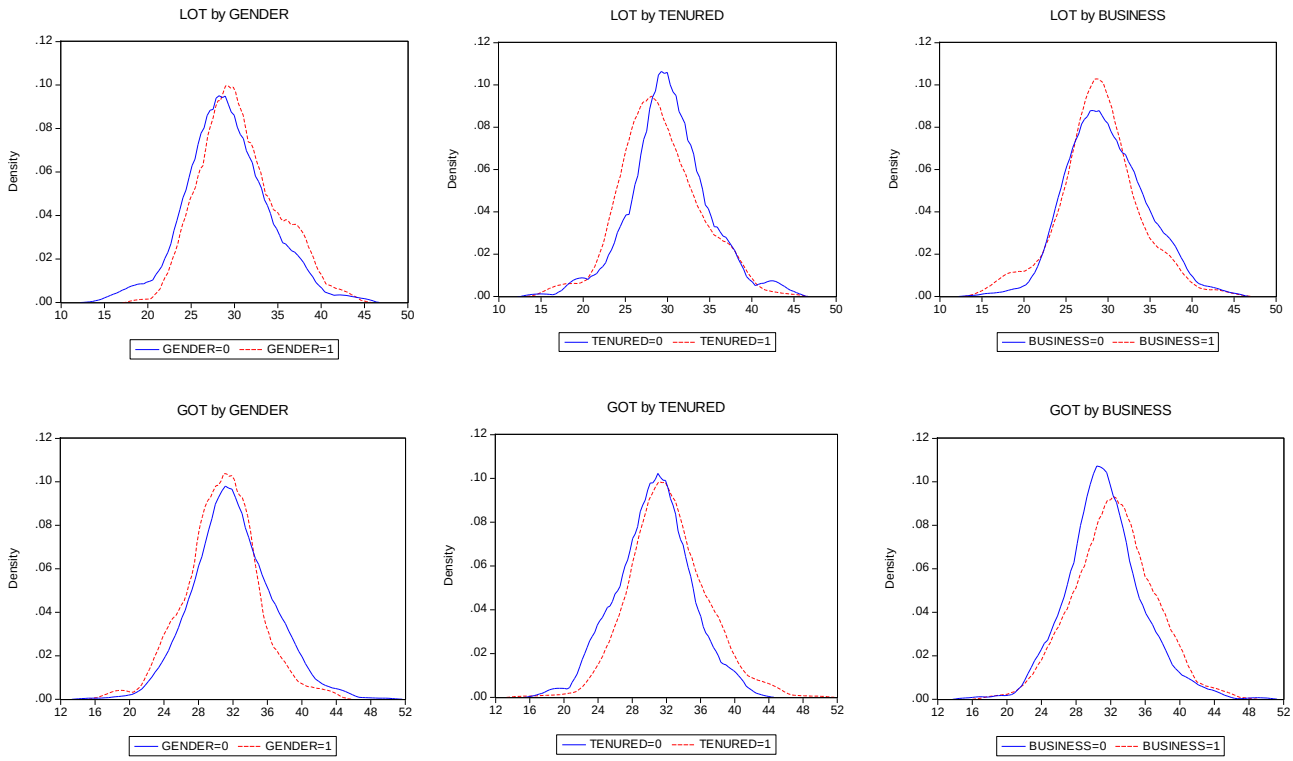
The views of economics faculty also seem to differ by certain individual and environmental characteristics. Tests for equality of means across groups indicate that gender, tenure status and business school affiliation are all correlated with learning and grade orientation. Table 3 presents summary results showing that females tend to be less GO and more LO compared to males; tenured faculty tend to be more GO and less LO compared to non-tenured faculty; and economics faculty affiliated with business schools tend to be more GO and less LO than economics faculty not affiliated with business schools. We also perform ordered logit regressions to account for correlation between these three factors. The results do not change. These regressions also show no differences in LO and GO based on number of course sections taught, number of students taught, or the course level (principles, intermediate, upper level or graduate). Complete results are shown in Appendix 2.

Table 3: Differences in learning orientation (LO) and grade orientation (GO) by individual and institutional characteristics

Measure	Differences by	df	t-value	Probability
GOT	Gender	770	3.56	0.0004
	Tenure Status	772	–5.19	0.0000
	Business School affiliation	762	–3.33	0.0009
LOT	Gender	770	–3.68	0.0003
	Tenure Status	772	3.17	0.0016
	Business School affiliation	762	2.55	0.0110

The differences by gender, tenure, and school affiliation can be seen in their distributions (kernel densities), as shown in Figure 1. We see a general shift of the distribution in each case. We hesitate to speculate on the reasons for the differences between males and females. With regard to differences by tenure status, one interpretation is that to get tenure, economics faculty are likely to have been judged hard graders (been very concerned with grades), based on our survey findings. In other words, faculty who appear to be soft graders may tend not to get tenure. Likewise, business schools may have an environment that encourages GO and less LO, or encourages faculty with those characteristics to select business schools. Given the limitations of our data, we leave for future research a rigorous examination of the reasons for these differences.

Figure 1: Kernel densities for learning orientation (LO) and grade orientation (GO) by gender (0 = males), tenure status (0 = non-tenured), business school affiliation (0 = non-business school economics faculty)



4. Free-form comments

Two hundred and two respondents (25 per cent of all respondents) submitted free-form comments. Many of the comments fall into three closely related areas: (1) standards and cross-faculty comparison of grade distributions; (2) the influence of grades on student evaluations and the influence of those evaluations on personnel decisions; and (3) grade inflation.

Standards and pressure to conform

Many respondents were concerned with pressure to conform to department grading norms. Views were wide-ranging. For example, it was often stated or implied that being at the low end of the distribution (relative to colleagues) is better than being at the high end. Others cited pressure from deans to provide higher grades and not to fail students.

Ten per cent of respondents indicated (on item #31) that department or college grading expectations had a 'great deal' of influence on how they grade. Another 33 per cent responded 'some' to that question. Roughly 29 per cent responded 'not at all'.

Many indicated that they adjust their grade distribution to be near the department average. 'I do compare my grade distributions with colleagues to determine if I am somewhat consistent.' 'I give fewer As, but aside from that I want my grade distributions to resemble those of my colleagues.' 'Implicitly I adjust the level of the course taught so the grade distribution looks like the rest of my

colleages [sic].' This last comment seems natural – standards must vary from school to school (introductory economics at Harvard is very different than at SUNY Oneonta). It also suggests a relative standard, with the measurement of student performance in one class being gauged by student performance in another.

One respondent wrote, 'In my intro classes it's difficult to hold higher standards than other faculty in my department. If I do, then about 30 per cent of my students hate me…' The same respondent continues, 'I have been redesigning my intro micro class for the past seven years and I believe that I am getting closer to an optimal model using Aplia, clickers and no exams.'

Several faculty compared the grades awarded in economics to those awarded in other disciplines. Lower grades in economics were most often seen as an indication of higher standards and rigour. They also may be driving students away from economics. As one put it, 'The signal value of grades is therefore eroded as many students do not really know where their comarative [sic] advantage lies.'

Not once did a respondent suggest that low grades are an indicator of poor teaching or irrelevant topics. Always, poor/low/harsh/tough grades were seen as upholding high standards. Almost always high standards were relative to other disciplines, which were often viewed as vaguely loose in their standards.

Grades, student evaluations, and personnel decisions

The role of course evaluations in tenure and promotion decisions was mentioned often. A few respondents explicitly stated that they did not grade as harshly as they might otherwise because harsh grades would hurt their student evaluations, which will then hurt their tenure, promotion and salary prospects. As one respondent noted, 'there is a noticable [sic] cost to holding to standards'. Another wrote, there is 'pressure not to set standards or expectations too high'.

On the other hand, too many high grades cause alarm. One respondent noted, 'when evaluating faculty colleagues exceptionally high student evaluations coupled with exceptionally high grades sets off alarms'. Another noted, 'Grading policies of faculty are more affected by the weight placed on course evaluations in promotion and tenure than any other policy.'

Several other comments indicated similar concern. The pernicious influence of grades in personnel decisions seems to be a cause for alarm. Students play a key role. One responder wrote, 'Students (most of them) are obsessed with grades. They will do anything to get a higher grade (again, not all students but most). They will try to move their grade up by tallying their points against my grade cuts and asking (or demanding) a point or two more.'

Grade inflation

The concerns with promotion and tenure and standards were intertwined with a concern with grade inflation. 'I want the grades in my class to be comparable to grades in other classes and therefore participate in the inflation' one respondent wrote.

While many economics faculty bemoan the pressure to accommodate students' expectations for high grades ('The pressure to give high grades both from the departmental colleagues and from the Dean of the College have reached epic levels'), others note pressure from colleagues and deans to avoid grade inflation ('Our department and college is very concerned about grade infation [sic]. If we started giving an unusual number of As, someone would speak to us about it to be sure we were upholding standards').

Grade inflation in other departments was seen by one faculty as an explanation for economics faculty not winning teaching awards ('It is however the case that faculty from our department almost never win college or university teaching awards, because these rely heavily on student's evaluations (which are biased upward when grades are inflated)').[10]

5. Conclusion

Grades are pervasive in higher education. Financial aid, scholarships, continuation in a major, parental approval, potential employment, and acceptance to graduate school are all tied to some extent to grades. It is only natural that faculty and students exhibit a 'grade orientation'. Indeed, economics faculty in our study agree that it is 'useful to use grades as incentives to increase student performance'. But, by emphasising grades as an incentive, economics faculty may be promoting the very orientation toward grades that many consider to be a problem. In our survey, 41 per cent of economic faculty agreed or strongly agreed that 'Students' concern about grades often interferes with learning in my classroom.' Certainly, faculty want students to be less concerned with grades and more focused on learning. Students apparently want this also, yet feel constrained by the emphasis placed on grades by teachers, parents, and others (Pollio and Beck, 2000).

Milton *et al.* (1986, p. 141) report that, 'Faculty may emphasise grades in their classrooms more than they need to or should. Faculty members have it within their power to reduce this pernicious and distorting aspect of educational practice that often seems to work against learning. If faculty would relax their emphasis on grades, this might serve not to lower standards but to encourage an orientation toward learning.' Based on our survey results, this is likely to be viewed with considerable skepticism by economics faculty. Yet, decades of empirical work in educational psychology, and even some recent findings in economics education, lend support to this claim.

There are practical issues limiting one's ability to de-emphasise grades. One respondent wrote: 'My approach to grading has to take into account the system of grades that students have learned in 14+ years of schooling. I would LOVE to do away with grades altogether, but the students in a single semester course do not often know how to act in that new environment...' The job is even more complicated when one considers the need also to convince colleagues, department chairs, deans and others. But note that grades do not have to be completely removed, simply de-emphasised. Repeated and widespread de-emphasis of grades in economics classes (as may already occur in other college classes) may be necessary to undo students' learned behaviour from many prior years.

For those looking for concrete recommendations for de-emphasising grades we provide the following shortlist, incorporating some findings from educational psychology, with examples of methods the authors have used in their classes:[11]

1. Use extrinsic rewards sparingly and in a non-controlling manner.
    a. Do not grade classroom experiments or class discussion, although other 'rewards' (e.g. small sweets) may be used without an emphasis on the reward or the performance itself. Appealing experiments and discussion will be enough to stimulate student effort. If students are not participating, change the activity (Dickie, 2006; Moeller and Reschke, 1993; Hahn *et al.*, 1989).

[10] A more detailed discussion of grade inflation is left for a future paper. But interested readers can examine a California State–Northridge study of grade inflation available at http://www.csun.edu/coc/report06.html. The website contains supporting data by school and department.

[11] Barbara Gross-Davis of UC-Berkeley provides details and additional suggestions, with references (http://teaching.berkeley.edu/bgd/motivate.html).

   b. Grade assignments satisfactory or unsatisfactory (with a minimum acceptable level of quality, e.g. B–, necessary for satisfactory work) as this will probably be seen as non-controlling, given that perfection is not required for full credit (Deci *et al.*, 1999; Grolnick and Ryan, 1987; Ames and Ames, 1991).

2. Provide choice (concerning assignments, topics to be discussed, due dates) (Anderman and Midgley, 1998).
   a. Allow students to choose which assignments to submit for credit (either requiring a minimum number of satisfactory assignments or reducing the weight on exam scores for each satisfactory assignment).
   b. Allow students to choose topics for class discussion from a list of possible topics.

3. Promote mastery of learning by providing opportunities to revise unsatisfactory work.
   a. Encourage students to revise and resubmit unsatisfactory work – even if it is graded 0-100 or A-F.

4. Avoid competition; base evaluation on criterion-referenced standards.
   a. Communicate to students that work is evaluated on set standards, not relative to other students' work (Urdan *et al.*, 1998; Thompson and Perry, 2005).

5. Encourage attributing success to effort and interpreting mistakes as learning opportunities.
   a. Repeatedly impress upon students that effort is the most important determinant of success in your course (not ability or luck); offer students concrete examples of how and on what they should be exerting effort. At the same time, make it clear to students that they must ultimately show a certain level of proficiency for a successful learning experience.
   b. Praise effort rather than outcome (Henderlong and Lepper, 2002).
   c. Allow students to replace poor grades (e.g. by allowing the score on a comprehensive final exam to replace lower exam scores), keeping them 'in the game' throughout the semester.

The healthy response rate to our survey is an indicator of the high interest in this topic among economics faculty. For many, grades have become a sore spot. This paper provides a record of the attitudes and behaviours of economics faculty on a variety of grade-related topics and serves as an important attempt to open a dialogue among economics faculty and between faculty of economics and other disciplines. Comparing the views of economics faculty to the empirical evidence sheds light on the likely effects of our common policies. Future research should apply insights from educational psychology to the economics classroom. This would extend beyond simple replication of prior studies – which may be warranted at first – to an extension, incorporating fundamental economic concepts such as opportunity costs and formal modelling (possibly of the type recently being explored in labour economics).[12] The focus of current research in economics education would then expand from content and methods of instruction, both very important areas, to include student motivation and incentives.

[12] Frey (1998), Falk and Kosfeld (2006) and Murdock (2002).

References

Allgood, S., Bosshardt, W., Van der Klaauw, W. and Watts, M. (2004). 'What Students Remember and Say about College Economics Years Later', *American Economic Review,* Vol. 94(2), pp. 259-265.

Ames, R. and Ames, C. (1991). 'Motivation and Effective Teaching', in (B. F. Jones and L. Idol, eds) *Educational Values and Cognitive Instruction: Implications for Reform,* Hillsdale, N. J.: Erlbaum.

Anderman, L. and Midgley, C. (1998). 'Motivation and middle school students', ERIC Clearinghouse on Elementary and Early Childhood Education. ERIC Document Reproduction Service No. ED 421 281.

AEA CSWEP (2008). 'Report of the Committee on the Status of Women in the Economics Profession', *American Economic Review,* Vol. 98(2), pp. 611-18.

Baird, L. (1985). 'Do grades and tests predict adult achievement?', *Research in Higher Education*, Vol. 23, pp. 3-85.

Becker, W. E. and Watts, M. (1996). 'Chalk and Talk: A National Survey on Teaching Undergraduate Economics', *American Economic Review,* Vol. 86(2), pp. 448-53.

Becker, W. E. and Watts, M. (2001a). 'Teaching Economics at the Start of the 21st Century: Still Chalk and Talk', *American Economic Review,* Vol. 91(2), pp. 446-51.

Becker, W. E. and Watts, M. (2001b). 'Teaching Methods in U.S. Undergraduate Courses', *Journal of Economic Education,* Vol. 32(3), pp. 269-79.

Betts, J. R. and Grogger, J. (2003). 'The Impact of Grading Standards on Student Achievement, Educational Attainment, and Entry-Level Earnings', *Economics of Education Review,* Vol. 22(4), pp. 343-52.

Bose, J. (2001). 'Nonresponse Bias Analyses At The National Center For Education Statistics', Proceedings of Statistics Canada Symposium. Accessed at http://www.fcsm.gov/committees/ihsng/StatsCan2_JB.pdf.

Butler, R. (1987). 'Task-Involving and Ego-Involving Properties of Evaluation: Effects of Different Feedback Conditions on Motivational Perceptions, Interest, and Performance', *Journal of Education Psychology,* Vol. 79(4), pp. 474-82.

Cherry, T. L. and Ellis, L. V. (2005). 'Does Rank-Order Grading Improve Student Performance? Evidence from a Classroom Experiment', *International Review of Economics Education,* Vol. 4(1), pp. 9-19.

Cohen, P. A. (1984). 'College grades and adult achievement: A research synthesis', *Research in Higher Education,* Vol. 20, pp. 281-94.

Davidson, R. C. and Lewis, E. L. (1997). 'Affirmative Action and Other Special Consideration Admissions at the University of California, Davis, School of Medicine', *Journal of the American Medical Association,* Vol. 278(14), pp. 1153-8.

Deci, E., Koestner, R. and Ryan, R. (1999). 'A Meta-analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation', *Psychological Bulletin,* Vol. 125(6), pp. 627-68.

Dickie, M. (2006). 'Experimenting: Does It Increase Learning in Introductory Microeconomics?', *Journal of Economic Education*, Vol. 37(3), pp. 267-88.

Eison, J., Janzow, F. and Pollio, H. R. (1993). 'Assessing Faculty Orientations Towards Grades and Learning: Some Initial Results', *Psychological Reports,* Vol. 73, pp. 643-56.

Falk, A. and Kosfeld, M. (2006). 'The Hidden Costs of Control', *American Economic Review,* Vol. 96(5), pp. 1611-30.

Frey, B. (1998). *Not just for the money: an economic theory of personal motivation,* Beacon Press: Boston.

Grolnick, W. and Ryan, R. (1987). 'Autonomy in Children's Learning An Experimental and Individual Difference Investigation', *Journal of Personality and Social Psychology,* Vol. 52(5), pp. 890-8.

Grove, W. A. and Wasserman, T. (2006). 'Incentives and Student Learning: A Natural Experiment with Economics Problem Sets', *American Economic Review,* vol. 96(2), pp. 447-52.

Hahn, S., Stassen, T. and Reschke, C. (1989). 'Grading Classroom Oral Activities: Effects on Motivation and Proficiency', *Foreign Language Annals,* Vol. 22, pp. 241-52.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M. and Elliot, A. J. (2000). 'Short-term and long-term consequences of achievement goals: Predicting interest and performance over time', *Journal of Educational Psychology,* Vol. 92, pp. 316-30.

Harter, S. (1978). 'Pleasure Derived from Challenge and the Effects of Receiving Grades on Children's Difficulty Level Choices', *Child Development,* Vol. 49, pp.788-99.

Henderlong, J. and Lepper, M. R. (2002). 'The Effects of Praise on Children's Intrinsic Motivation: A Review and Synthesis', *Psychological Bulletin,* Vol. 128(5), pp. 774-95.

Janzow, F. and Eison, J. (1990). 'Grades: Their Influence on Students and Faculty', *New Directions for Teaching and Learning*, Vol. 42(summer), pp. 93-102.

Kamath, R. R., Thomas, E. G. and Wang, Y. (2007). 'Academic and personal characteristics of economics faculty in U.S. colleges and universities', *International Journal of Business Research*, Vol. 7(3), pp.29-40.

Karabenick, S. (2004). 'Perceived Achievement Goal Structure and College Student Help Seeking', *Journal of Educational Psychology*, Vol. 96(3), pp. 569-81.

Kirschenbaum, H., Napier, R. and Simon, S. (1971). *Wad-Ja-Get? The Grading Game in American Education*, New York: Hart.

Linnenbrink, E. (2005). 'The Dilemma of Performance-Approach Goals: The Use of Multiple Goal Contexts to Promote Students' Motivation and Learning', *Journal of Educational Psychology*, Vol. 97(2), pp. 197-213.

Meece, J., Anderman, E. and Anderman. L. (2006). 'Classroom Goal Structure, Student Motivations, and Academic Achievement', *Annual Review of Psychology*, Vol. 57, pp. 487-503.

Milton, O., Pollio, H. and Eison, J. (1986). *Making Sense of College Grades,* San Francisco: Jossey-Bass.

Moeller, A. J. and Reschke, C. (1993). 'A Second Look at Grading and Classroom Performance: Report of a Research Study', *The Modern Language Journal*, Vol. 77(2), pp. 163-69.

Murdock, K. (2002). 'Intrinsic motivation and optimal incentive contracts', *RAND Journal of Economics*, Vol. 33(4), pp. 650-671.

Oppenheim, A. N. (1966). *Questionnaire Design and Attitude Measurement,* New York: Basic Books.

Pekrun, R., Elliott, A. and Maier, M. (2006). 'Achievement Goals and Discrete Achievement Emotions: A Theoretical Model and Prospective Test', *Journal of Educational Psychology*, Vol. 98(3): 583-97.

Pollio, H. R. and Beck, H. P. (2000). 'When the Tail Wags the Dog: Perceptions of Learning and Grade Orientation in, and by, Contemporary College Students and Faculty', *The Journal of Higher Education*, Vol. 71(1), pp. 84-102.

Schaur, G., Watts, M. and Becker, W. E. (2008). 'Assessment practices and Trends in Undergraduate Economics Courses', *American Economic Review*, Vol. 98(2), pp. 552-56.

Senko, C. and Harackiewicz, J. M. (2005). 'Regulation of Achievement Goals: The Role of Competence Feedback', *Journal of Educational Psychology*, Vol. 97(3), pp. 320-36.

Skinner, B. F. (1976). *About Behaviorism,* Vintage: New York.

Thompson, T. (1994). 'Self-worth protection: Review and implications for the classroom', *Educational Review,* Vol. 46(3), pp. 259-74.

Thompson, T. and Perry, Z. (2005). 'Is the Poor Performance of Self-Worth Protective Students Linked with Social Comparison Goals?', *Educational Psychology*, Vol. 25(5), pp. 471-90.

Urdan, T., Midgley, C. and Anderman, E. (1998). 'The Role of Classroom Goal Structure in Students' Use of Self-Handicapping Strategies', *American Educational Research Journal,* Vol. 35(1), pp. 101-22.

Walstad, W. B. (2006). 'Assessment of Student Learning in Economics', in (W. E. Becker, M. Watts and S. R. Becker, eds) *Teaching Economics, More Alternatives to Chalk and Talk,* Cheltenham, UK: Edward Elgar.

Appendix 1: Demographic questions

21. Are you Male or Female?
    ___ Male ___ Female

22. What is your academic status (rank)?
    ___ Assistant Professor (tenure track)
    ___ Associate Professor (tenured)
    ___ Professor (tenured)
    ___ Full-time: Lecturer/Visiting (non-tenured)
    ___ Part-time: Lecturer/Visiting (non-tenured)
    ___ Other, please specify

23. How many years have you been teaching? ___

24. What is the highest degree awarded by your department?
    ___ Bachelors
    ___ Masters
    ___ Doctorate
    ___ Other, please specify

25. In which school/division is your department?
    ___ School of Business
    ___ School of Social Sciences
    ___ Other, please specify

26. Approximately how many full-time economics faculty are in your department? ___

27. How many course sections and students do you teach in a typical academic year in each of the areas listed below? (Leave blank areas not taught.)

	Course sections	Number of students (total)
Principles, undergrad	_____	_____
Intermediate, undergrad	_____	_____
Upper level, undergrad	_____	_____
MA or MBA	_____	_____
PhD	_____	_____

28. For each course level that you teach, indicate the TWO most significant evaluation tools by placing a 1 in the box for the method that accounts for the greatest portion of the course grade and a 2 in the appropriate box for the second greatest portion. Leave blank if you do not teach the course level shown.

	Essay / Short answer exams	Multiple choice exams	Homework	Quizzes	Papers	Class participation/ presentations
Introductory Level						
Intermediate						
Upper level, graduate						

29. What subject area do you regularly teach? Check all that apply.
    Select no more than 5.
    ___ Micro
    ___ Macro
    ___ International (trade/finance)
    ___ Econometrics/statistics
    ___ Financial (incl. Money & Banking)
    ___ Public Finance
    ___ Urban/Regional
    ___ Industrial Org/Regulation/Government
    ___ Labor
    ___ Environmental
    ___ Other, please specify

30. What percentage of all the students that you teach in a typical academic year are of each type listed below? (Use 100 for 100%, 50 for 50%,....). Skip this question if you are unsure.

	Percentage
Econ majors	_____
Business / Business economics majors (non-econ)	_____
Non-econ, non-business	_____

31. To what extent does department or college grading expectations, whether explicit or implicit, influence how you grade students or the course grades you submit? We invite you to use the comment box at the end of the survey to explain.
    ___ Not at All
    ___ Very Little
    ___ Some
    ___ A great Deal

32. Please enter comments in the space below.

Appendix 2: Complete results

TABLE A-1: Ordered Logit Regression with Learning Oriented (LO) total score as the dependent variable, n = 744.

Variable	Coefficient	Std. Error	z-Statistic	Prob.
GENDER	0.4304	0.15128	2.845	0.0044
TENURED	–0.5530	0.14034	–3.940	0.0001
BUSINESS	–0.3034	0.13906	–2.182	0.0291
PRINCIPLES_S	–0.0011	0.00049	–2.260	0.0238
INTERMEDIATE_S	0.0011	0.00127	0.935	0.3495
UPPER_S	–0.0024	0.00180	–1.385	0.1661
GRAD_S	0.0001	0.00249	0.079	0.9369
PRIN_SECT	0.0023	0.03511	0.066	0.9473
INTM_SECT	–0.0853	0.07445	–1.147	0.2514
UPPER_SECT	0.0551	0.07305	0.755	0.4501
GRAD_SECT	0.0456	0.08651	0.528	0.5974
	Limit Points			
LIMIT_16:C(12)	–7.205937	1.013602	–7.109238	0.0000
LIMIT_17:C(13)	–6.102603	0.600611	–10.16066	0.0000
LIMIT_18:C(14)	–5.247706	0.412585	–12.71908	0.0000
LIMIT_19:C(15)	–4.885428	0.356836	–13.69096	0.0000
LIMIT_20:C(16)	–4.404930	0.299688	–14.69840	0.0000
LIMIT_21:C(17)	–4.030918	0.266116	–15.14720	0.0000
LIMIT_22:C(18)	–3.864262	0.253745	–15.22893	0.0000
LIMIT_23:C(19)	–3.591904	0.236528	–15.18594	0.0000
LIMIT_24:C(20)	–3.008977	0.208478	–14.43305	0.0000
LIMIT_25:C(21)	–2.548581	0.193012	–13.20430	0.0000
LIMIT_26:C(22)	–2.041067	0.181379	–11.25305	0.0000
LIMIT_27:C(23)	–1.650002	0.175305	–9.412186	0.0000
LIMIT_28:C(24)	–1.199075	0.170122	–7.048335	0.0000
LIMIT_29:C(25)	–0.758983	0.166493	–4.558650	0.0000
LIMIT_30:C(26)	–0.348498	0.164692	–2.116053	0.0343
LIMIT_31:C(27)	0.031286	0.164313	0.190403	0.8490
LIMIT_32:C(28)	0.341794	0.165026	2.071147	0.0383
LIMIT_33:C(29)	0.661109	0.166923	3.960572	0.0001
LIMIT_34:C(30)	1.039585	0.171133	6.074714	0.0000
LIMIT_35:C(31)	1.409345	0.177819	7.925740	0.0000
LIMIT_36:C(32)	1.608473	0.182752	8.801389	0.0000
LIMIT_37:C(33)	1.943836	0.193317	10.05515	0.0000
LIMIT_38:C(34)	2.409252	0.213821	11.26762	0.0000
LIMIT_39:C(35)	3.061541	0.258240	11.85539	0.0000
LIMIT_40:C(36)	3.389590	0.289315	11.71592	0.0000
LIMIT_41:C(37)	3.602639	0.313196	11.50284	0.0000
LIMIT_42:C(38)	3.978397	0.363625	10.94094	0.0000
LIMIT_43:C(39)	4.574054	0.470187	9.728159	0.0000
LIMIT_44:C(40)	4.798743	0.520638	9.217035	0.0000

TABLE A–2: Ordered Logit Regression with Grade Oriented (GO) total score as the dependent variable, n = 744.

Variable	Coefficient	Std. Error	z–Statistic	Prob.
GENDER	–0.3657	0.15037	–2.432	0.0150
TENURED	0.7170	0.14381	4.986	0.0000
BUSINESS	0.5251	0.14216	3.693	0.0002
PRINCIPLES_S	1.21E–05	0.00038	0.031	0.9745
INTERMEDIATE_S	–0.0004	0.00130	–0.366	0.7137
UPPER_S	0.0008	0.00186	0.433	0.6645
GRAD_S	–0.0002	0.00312	–0.093	0.9253
PRIN_SECT	0.0325	0.03247	1.003	0.3155
INTM_SECT	0.0473	0.07492	0.631	0.5277
UPPER_SECT	–0.0143	0.06775	–0.212	0.8317
GRAD_SECT	–0.0683	0.09543	–0.716	0.4736
	Limit Points			
LIMIT_18:C(12)	–6.097132	1.009869	–6.037548	0.0000
LIMIT_19:C(13)	–5.401951	0.721012	–7.492174	0.0000
LIMIT_20:C(14)	–4.994411	0.594322	–8.403541	0.0000
LIMIT_22:C(15)	–4.295928	0.431985	–9.944622	0.0000
LIMIT_23:C(16)	–3.885608	0.362085	–10.73119	0.0000
LIMIT_24:C(17)	–3.014938	0.260491	–11.57405	0.0000
LIMIT_25:C(18)	–2.333114	0.213536	–10.92611	0.0000
LIMIT_26:C(19)	–1.948111	0.195835	–9.947739	0.0000
LIMIT_27:C(20)	–1.526328	0.182105	–8.381567	0.0000
LIMIT_28:C(21)	–1.061226	0.172696	–6.145068	0.0000
LIMIT_29:C(22)	–0.697445	0.168202	–4.146467	0.0000
LIMIT_30:C(23)	–0.271592	0.165295	–1.643076	0.1004
LIMIT_31:C(24)	0.190916	0.164249	1.162352	0.2451
LIMIT_32:C(25)	0.606880	0.165171	3.674253	0.0002
LIMIT_33:C(26)	1.044430	0.168113	6.212657	0.0000
LIMIT_34:C(27)	1.464461	0.172153	8.506735	0.0000
LIMIT_35:C(28)	1.881892	0.177475	10.60370	0.0000
LIMIT_36:C(29)	2.210563	0.183016	12.07852	0.0000
LIMIT_37:C(30)	2.538226	0.190195	13.34538	0.0000
LIMIT_38:C(31)	2.931805	0.201160	14.57449	0.0000
LIMIT_39:C(32)	3.315835	0.215143	15.41220	0.0000
LIMIT_40:C(33)	3.848507	0.242929	15.84212	0.0000
LIMIT_41:C(34)	4.253969	0.272732	15.59759	0.0000
LIMIT_42:C(35)	4.534038	0.298792	15.17457	0.0000
LIMIT_43:C(36)	4.828431	0.331857	14.54975	0.0000
LIMIT_44:C(37)	5.241158	0.389638	13.45135	0.0000
LIMIT_45:C(38)	6.231469	0.600150	10.38318	0.0000
LIMIT_46:C(39)	6.639020	0.725849	9.146554	0.0000
LIMIT_49:C(40)	7.334229	1.013344	7.237650	0.0000

Contact details

Lester Hadsell
Department of Economics
State University of New York, College at Oneonta
Oneonta, NY 13820, USA
Email : hadsell@oneonta.edu
Tel: 607-436-2448

Raymond MacDermott
331 Scott Shipp Hall
Virginia Military Institute
Lexington, VA 24450, USA
Email: macdermottrj@vmi.edu
Tel : 540-464-7405

The Road to Transfer: Concept and Context Approaches to the Subject of Economics in Secondary School

Lenie Kneppers, Carla van Boxtel and Bernadette van Hout-Wolters

Abstract

In this study we investigated the effects of two forms of instruction: strengthening concepts and strengthening the ability to connect context with concepts. Although students may have acquired a reasonable amount of conceptual knowledge as a result of economics courses, two obstacles may prevent them from achieving transfer. One obstacle is a lack of a rich conceptual network; another is the inability to make connections between the conceptual network and realistic problems. The aim of this study was to find out what contributes most to the ability to transfer: strengthening conceptual knowledge *or* strengthening the making of connections. Some 139 students of the pre-final year of pre-university education participated in an experiment with two conditions and with a pre-test and a post-test. All students performed significantly better on the post-test in which conceptual knowledge was measured compared to the pre-test. No significant differences were measured between the two instructions on the post-test on transfer. We concluded that making connections is a difficult skill for students to learn. Context directed instruction definitively supports student's knowledge of concepts.

JEL classification: A20, A21, B40

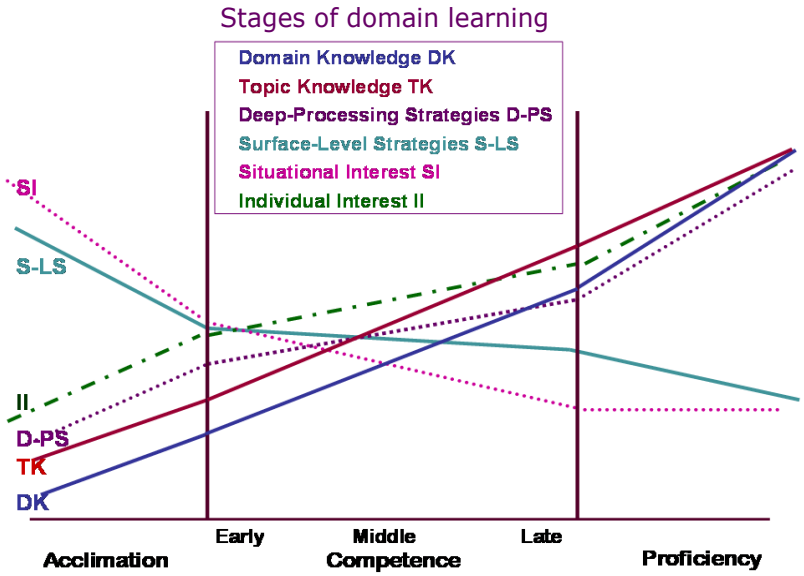
1. Introduction

The aim of economics education in secondary schools is shifting from a predominantly academic approach towards more emphasis on teaching that meets the needs of students in their current and future lives. From this perspective, education in economics should enable students to use economic knowledge in daily life, even in contexts that are different from those in which they originally acquired this knowledge.

When students can apply acquired knowledge in a more or less novel situation, they have achieved transfer (Marini and Genereux, 1995). Stark *et al.* (1999) concluded that it can be useful to distinguish middle transfer as well as near transfer and far transfer. We consider three transfer stages in this study but we use the terms: near transfer, semi-far transfer and far transfer. We are talking about near transfer if concepts and context are the same as in the instruction; semi-far transfer if concepts are the same but the context is different; and far transfer if both concepts and context are different from (but related to) those in the instruction setting. We can explain this with an example.

In her model of domain learning, Alexander (2003) considers transfer as an ability of experts. She describes the long road to expertise in three stages: *acclimation*, *competence* (*early, middle* and *late*) and *proficiency*. Each of these stages is characterised by a certain amount of domain knowledge, interest and types of strategies used. Figure 1 shows how the development works in the three stages.

Figure 1: Knowledge interest and use of strategies in different stages (Alexander, 2005)



Students in secondary education are beginners in economic studies. They belong to the first stage: *acclimation*. They do not have much individual interest in the domain of economics (II), but what they do possess is situational interest (SI). Entering the second stage – *early competence* – individual interest is increasing as well as domain knowledge (DK) and knowledge of economic topics (TK).

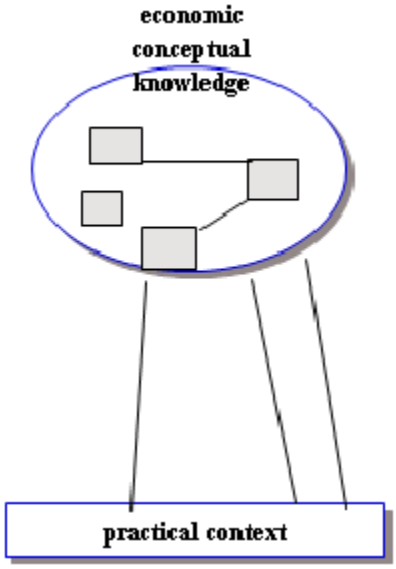
In secondary education students often start with learning concepts. Although in the upper levels of secondary education students may have acquired a reasonable amount of conceptual knowledge as a result of economic courses, two obstacles may prevent them from achieving transfer. One possible hindrance is the lack of a rich and coherent conceptual network, whereas deeply understood and well organised domain knowledge is viewed as a prerequisite for achieving transfer (Mayer, 2004; Sternberg, 2003). The conceptual network attained by students after instruction often shows gaps. Concepts are missing or relations between the concepts are not well defined. This kind of obstacle is normally the effect of students studying for a knowledge-oriented achievement test. They study with the aim of reproducing knowledge in the exam but not with the aim of thoroughly understanding it.

It may be sufficient for near transfer if the context is very similar to the studied problem, but it will be failing for far transfer where the context is different from the studied problem. An example of near transfer is when the concept of risk aversion is used in the context of assurance in both construction and the transfer task. Transfer is semi-far when in the instruction the concept of risk aversion is used in the context of assurance, whereas in the transfer task the concept of risk aversion is used in the context of investments. Far transfer occurs when in the instruction the concept of risk aversion is used in the context of assurance, whereas in the transfer task the concept of own risk (related to risk aversion) is used in the context of a venture.

A second possible obstacle for achieving transfer is that students are hardly able to make connections between the acquired conceptual network and realistic social problems that can be looked at from an economic perspective. Conversely, they are unable to connect real life situations they encounter with the appropriate conceptual network they have in mind, that matches with the practical phenomena (Becker, 2004; Hansen *et al.*, 2002).

The two kinds of obstacles are visualised in Figure 2.

Figure 2: The two kinds of deficits that may prevent students from achieving transfer



In the upper part of this figure the oval represents the conceptual knowledge of complex domain knowledge. Complex domain knowledge exists if a network of concepts relate to one another. Bransford *et al.*, (1999) describe *conceptual networks* as 'knowledge organised around important ideas or concepts of a domain'.

The squares in the oval represent the economic concepts and the interconnecting lines represent the relationships between the concepts. The oval represents the larger economic concept to which these concepts belong. The number of squares is small and some squares in the oval are interconnected, whereas others are not. This indicates that the conceptual knowledge is relatively poor for a complex concept. As discussed above, we suppose that students have acquired a collection of concepts (with the aim of passing economic tests at the end of the course) rather than a solid conceptual network preparing them for approaching and analysing social problems in an economic way. In the lower part of Figure 2 the ability to make connections between the conceptual knowledge and practical contexts is represented by the lines linking the oval to the practical context. The small number of links in the figure indicates that we suppose students can make only a few links between conceptual knowledge and practical contexts that might be approached from an economic perspective.

2. Learning for transfer in school

We have noticed the two different obstacles for achieving transfer. The ability to transfer, however, is a necessity for students looking at real life problems as an economist would. Perkins (1992) wrote:

> 'Consequently, the ends of education are not achieved unless transfer occurs. Transfer is all the more important in that it can not be taken for granted. Abundant evidence shows that very often the hoped-for transfer from learning experiences does not occur… Thus the prospects and conditions of transfer are crucial educational issues.'[1]

[1] Perkins, D. N., 1992, p.3

The lack of a rich and coherent conceptual network of knowledge is a condition for transfer, but it does not lead to transfer in itself. The ability to connect context with concepts is another condition, but it does not lead to transfer in itself either, because besides the knowledge of the context, the student must possess conceptual knowledge to be able to connect the two. Both conditions are necessary.

The question can be asked: what contributes most to the ability of transfer when students have learnt the concepts in their economics course before – strengthening their conceptual knowledge (the concept road) or strengthening their ability to connect context with concepts (the context road)? The first approach corresponds with Bloom's taxonomy (1968). The original taxonomy consists of a cumulative hierarchy of three processes necessary to reach meaningful learning and transfer of knowledge: memorising, comprehending and applying in a context. Mayer (2002), Krathwohl (2002) and Anderson (2002) decided not to consider the hierarchy: students can come to an understanding of a concept *while* applying it in a context.

The question then arises as to what contributes most to the ability of transfer: strengthening conceptual knowledge *or* strengthening the making of connections between contexts and concepts? In other words, are we following the concept road or the context road to transfer?

The concept road to transfer

In literature about transfer, for example Gelman and Greeno (1989),and Salomon and Perkins (1989), there is an agreement that a basis of deeply understood or processed knowledge is a condition for transfer. Such a body of knowledge requires meaningful learning. Novak (2002) stated that meaningful learning is defined by a conscious choice of integrating new knowledge into existed knowledge. Existing mental models, however, can hinder the forming of the conceptual models required for the domain knowledge. This is the case when students possess conceptions that differ from formal domain conceptions. These alternative conceptions or misconceptions are well known in the science domain, but they also exist in economics. In the case of misconceptions it is necessary when acquiring formal domain knowledge that *conceptual change* takes place. Chinn and Brewer (1993) and Vosniadou (1994) described conceptual change as a process of accretion, refining, constructing and reconstructing of mental models. Students must have sufficient time to construct and reconstruct their mental models to develop a *well organised body of knowledge* to achieve transfer. In the concept road to transfer, the main focus is on the economic concepts and emphasis is put on the development of a rich and coherent network of related concepts. The idea is that such a network enables transfer.

The context road to transfer

The context road originated from a theory of learning, mostly referred to as 'situated learning' but also as 'situated cognition' or 'situated action' (Brown *et al.* 1989; Engeström *et al.*, 1999). This theory emphasises that all knowledge is situated in a particular time and place, and is a part of a culture in which the knowledge has been developed and used. This has many consequences for learning at school. The context at school is very different from that in real life, so it is not surprising that students do not connect economic concepts to practical contexts. They possess mainly inert knowledge that will sink in for no longer than is necessary for them to sit exams. They hardly ever get the opportunity to see the world as a historian (Whitehead, 1957) or, in our case, as an economist. That is why the out-of-school situation in which students function must be understood well:

> 'This is not to suggest that all students of math or history must be expected to become professional mathematicians or historians, but to claim that in order to learn these subjects (and not just learn about them) students need much more than abstract concepts and self-contained examples. They need to be exposed to the use of a domain's conceptual tools in authentic activity to teachers acting as

> practitioners and using these tools in wrestling with problems of the world. Such activity can tease out the way a mathematician or historian looks at the world and solves emergent problems.'[2]

Theories of situated cognition emphasise that students need to be exposed to the use of domain specific concepts, methods in authentic activity and to 'problems of the world' (e.g. Brown *et al.*, 1989; Gorodetsky *et al.*, 2003). Where it is not possible to join a community out of school for situated learning, 'the world' has to be brought into the school. School has to become a 'community of learners', where students are able to struggle with the problems of the world.

By context we understand (practical) situations, concerning economic phenomenon that are recognisable for students. These are recognisable, authentic situations in the roles that students have in their daily life and on which they have to make decisions. The starting point is that they form a well-organised knowledge base while working on contexts. Students reach deep understanding of concepts *while* using those concepts in a context, because they need these concepts to analyse the context problem.

This study

In this experimental study, we compare the effects of a concept road to transfer – strengthening/ reconstruction of concepts – and a context road to transfer – making connections between the authentic, practical context and concepts. Each of these approaches focuses on one of the obstacles to transfer described above. In both types of instruction it is possible to make use of concept mapping as a learning tool. Later in this article we go deeper into concept mapping as an appropriate learning activity for both instructions.

3. Research question

Above we mentioned that meaningful learning – compared to root learning – occurs when students actively construct, reconstruct and apply the mental models they possess. The formation of a well-organised body of knowledge which is the result of that activity is expected to have a positive influence on the ability for transfer. Conceptual knowledge is built in a flexible way and can be applied in different situations to successfully solve problems (Mayer, 2002; Novak, 2002; Perkins, 1992). The concept road to transfer is characterised by the acquisition of this well-organised knowledge *before* applying it in a context. Above we mentioned that learning must take place in contexts where meaning has been given to concepts. The context road to transfer is characterised by the acquisition of this well-organised knowledge *while* applying it in an everyday context.

The research question of our study is: *Which instruction, added to the usual lessons of economics in the fifth grade of pre-university education, is more effective in obtaining transfer: an instruction aimed at strengthening knowledge of concepts and relations between concepts, or an instruction based on constructing relations between knowledge of economic concepts and daily reality?*

[2] Brown, A. L., Collins, A. and Duguid, P. (1989), p. 36

4. Method

Design

This study is an experiment with two experimental conditions, the concept-condition and the context-condition, with a pre-test and a post-test. The pre-test was administered to check that groups did not differ in prior economic knowledge. There is one independent variable: the form of instruction. The dependent variables are near transfer, semi-far transfer and far transfer.

Participants

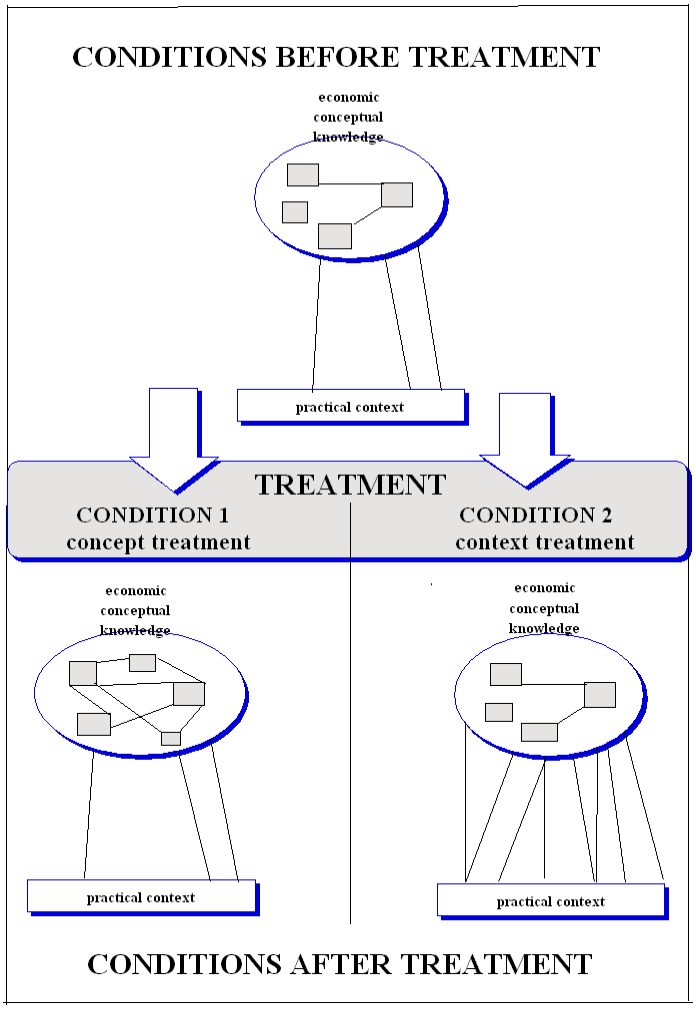
Some 139 students between the ages of 16 and 18, from eight classes of the fifth grade of pre-university education, and from six different schools, participated in our study. They had all chosen the three-year economics course. This research took place in the second year of the course. In all selected classes the economic subjects *The economic climate* and *Money circulation* – used in the research – were taught in the first year of the economics course. Links between these subjects had not been discussed so far. Because the concepts were not new to the students, we considered a two-hour-long instruction sufficient. Within each class, students were randomly assigned to one of the two conditions.

Instruction

The instruction for the concept condition consists of tasks which challenge students to reconstruct and consolidate their conceptual networks. The goal of this is that the first obstacle for transfer we mentioned – not enough concepts, wrongly understood concepts and incorrect relations between the concepts (see Figure 2) – is obviated completely or partially. The instruction for the context condition consists of tasks which challenge students to strengthen the connections between the conceptual framework and contexts, i.e. economic phenomena in the real world. This instruction corresponds with the second obstacle for transfer (see Figure 2).

Figure 3 below gives a schematic view of the aimed knowledge that is being acquired in the two conditions. The point of departure (in the upper part of the figure) is the same for the two conditions: the students have concept knowledge at their disposal (acquired in the school instruction), but the concept network is not complete and not all concepts are related to each other. Few connections can be seen between the conceptual network and the practical context. In the lower part of the figure our expectations after the instruction can be seen. The connections from the conceptual knowledge to the practical context are shown in the lines between the concepts and the practical situations. In the upper part of Figure 3 there are not many lines. The concept condition (bottom left) improves the conceptual network, and the context condition (bottom right) increases the number of connections between concept and practical context. It is possible that in the context condition, the conceptual network is also improved, since students will link these concepts to a concrete context, but we do not think students' conceptual network will reach the same quality as in the concept condition. We expected that an especially rich and coherent conceptual network would enable far transfer (see also our hypotheses below).

Figure 3: The desired changes in knowledge in concept and context conditions



As mentioned, before the experiment, in their normal school programme, the students followed a course on money circulation and the economic climate. In the pre-test, students' knowledge about these concepts was measured. In both conditions students were constructing concept maps and instruction session took two lessons of 50 minutes each. The instructions and the test were part of the normal schedule at school. The total research period covered 2.5 weeks, five lessons altogether: two for instructions and three for pre- and post-tests (measures of near, semi-far and far transfer).

The learning materials: concept mapping tasks and information paper about the economic climate and money circulation

For both conditions we developed a concept mapping task on which students worked in pairs. Concept mapping is a learning activity that can be used in both the concept road and the context road (see Appendices 1 and 2 for the full instructions). Novak (1990) describes a concept map as a compilation of concepts connected by relations. O'Donell *et al.* (2002) use the word 'knowledge map' rather than 'concept map'. Besides concepts, they say that a knowledge map can include dynamic relations, static relations and elaborative relations that contain information. A concept map can contain relations between concepts and practical situations. It can also contain prior knowledge which is important to activate (Beyerbach and Smith, 1990) as well as conditional and situational knowledge (Alexander, 2006). When students are co-constructing a concept map, the interaction is strengthening the benefit

of concept mapping. The collaboration with peers is motivating students to (re)construct their conceptual network (Roth and Roychoudhury, 1992).

An example of a map in the concept condition is shown in Figure 4 and that in the context condition In Figure 5.

Figure 4: An example of a map in the concept condition

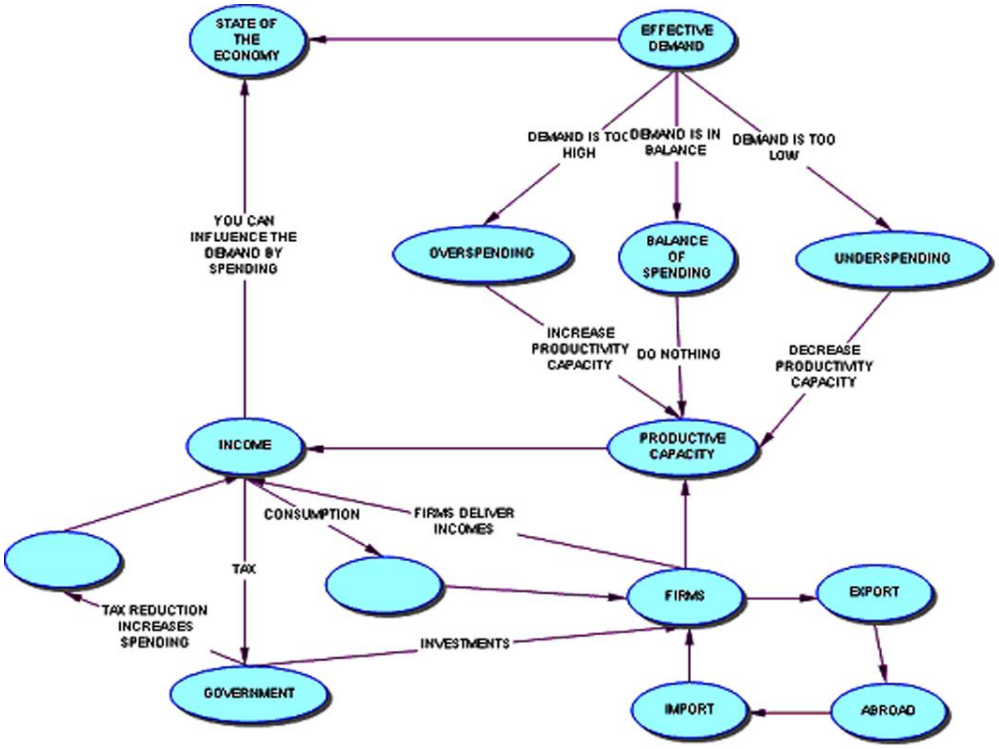
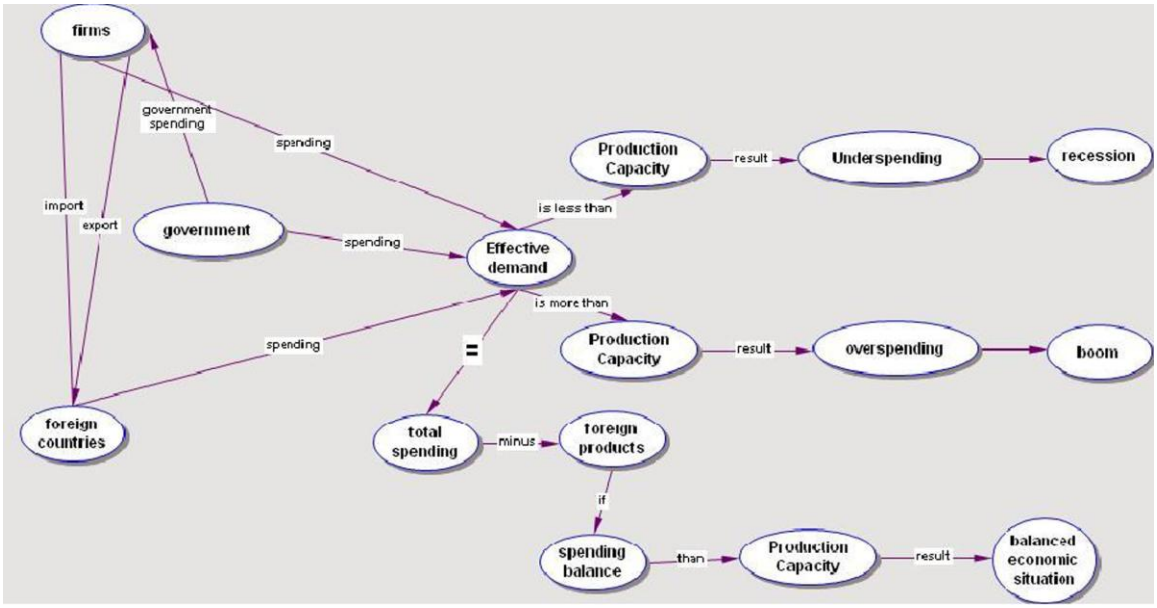


Figure 5: An example of a map in the context condition



Students did not have experience with concept mapping. They worked in the first instruction lesson individually on a concept map using paper, Post-its and pencils. After that they worked in pairs, constructing one concept map on the computer (using a program called Inspiration, www.inspiration.com). We decided to form pairs to promote interaction. Because students had to construct a joint concept map, it was desirable to work with not more than two students. The pairs were composed by using the middle group method (Pijls *et al.*, 2003). The pairs had to have discussions to reach a consensus. Students were asked to write down the questions they encountered while working on the concept map on a form. In the second instruction lesson students received information on where they could find the answers to their questions. With this information they had to reconstruct their first jointly made concept map. The formulation of the questions was a stimulus for discussion while constructing the concept map and for searching for information.

In the task for the concept condition, students were asked to make a scheme of labels and links to show how money circulation is related to the state of the economy in a year, when the state of the economy is in balance. Students were given three concepts (labels) they had to use and they were encouraged to use concepts they knew from earlier instruction. The task for the context condition was oriented towards a complex social problem. In the Netherlands a supplementary duty (25 cents) was set by the government on the price of petrol to discourage motorised traffic. Some members of Parliament wanted to discard this supplement retrospectively. Students were to discuss the possible economic effects of this proposal for society. Thus, they had to apply conceptual knowledge to a concrete economic problem. Students received some examples of practical outcomes and the same three concepts and were asked to make a concept map. They were explicitly stimulated to think from the practical context to economical concepts and *vice versa*.

The information used in this experiment differed from the textbooks students used in class. Besides the economic concepts, there was a lot of practical information, so students had an example of how to use economic concepts in practical contexts.

Tests

All students took four tests:

(a) a *pre-test* to measure prior domain specific conceptual knowledge. This test consisted of 25 questions. The test was constructed using a computer test program ('Wintoets'). The minimum score was 0, the maximum score was 72. The item homogeneity was not very high, but we considered it as just sufficient: Cronbach's alpha was 0.60. Because it was considered a test with divergent questions, the construct we measured was heterogeneous in nature; that is influencing the alpha (De Vellis, 1991). We used the computer program 'Wintoets' (www.drp.nl) for scoring the closed questions. The open questions were scored using a correction scheme.

(b) a *far transfer test* to measure students' ability to connect *new* knowledge to existing knowledge in a *new* context. The test was labelled '*Five circumstances in difficult economic times'* (see Appendix 3 for the full task). It entailed writing five texts associated with five economic cartoons drawn for a newspaper. The students were informed that the editor accidentally threw away the original stories associated with the cartoons. The students' task was to write appropriate economic texts for the five cartoons. The texts were scored on the

amount of connections between *T* (conText) and *C* (conCept) – the context scoring – and on the economic correctness of the *C* pointed units – the concept scoring. The minimum score was 0 and there was no maximum. The definitive concept scoring was determined by dividing the correct *C* units by the total *C* units. The inter-judge reliability for the context score was good (Cohen's kappa 0.85) and reliability for the concept score was sufficient (Cohen's kappa 0.64). When laying down the codes it was first established how to judge for a *C* (conCept) or a *T* (conText). The texts were scored by two researchers. Researcher 1 divided the texts into scoring units. Both researchers scored the *C* units on *C* or *T* and on economic correctness (see the scoring method in Appendix 4). The validity is partly guaranteed by the fact that both reviewers were economists so they could determine adequately when the concepts were used in an economic way and when they were not.

(c) a *near-transfer/semi-far transfer test*. This test consists of two parts:

- a concept post-test to measure conceptual domain knowledge after the experiment (a near transfer test for the concept condition and a semi-far transfer test for the context condition). This test is the same as the pre-test. The item homogeneity of this post-test (Cronbach's alpha 0.72) was higher than that of the pre-test.

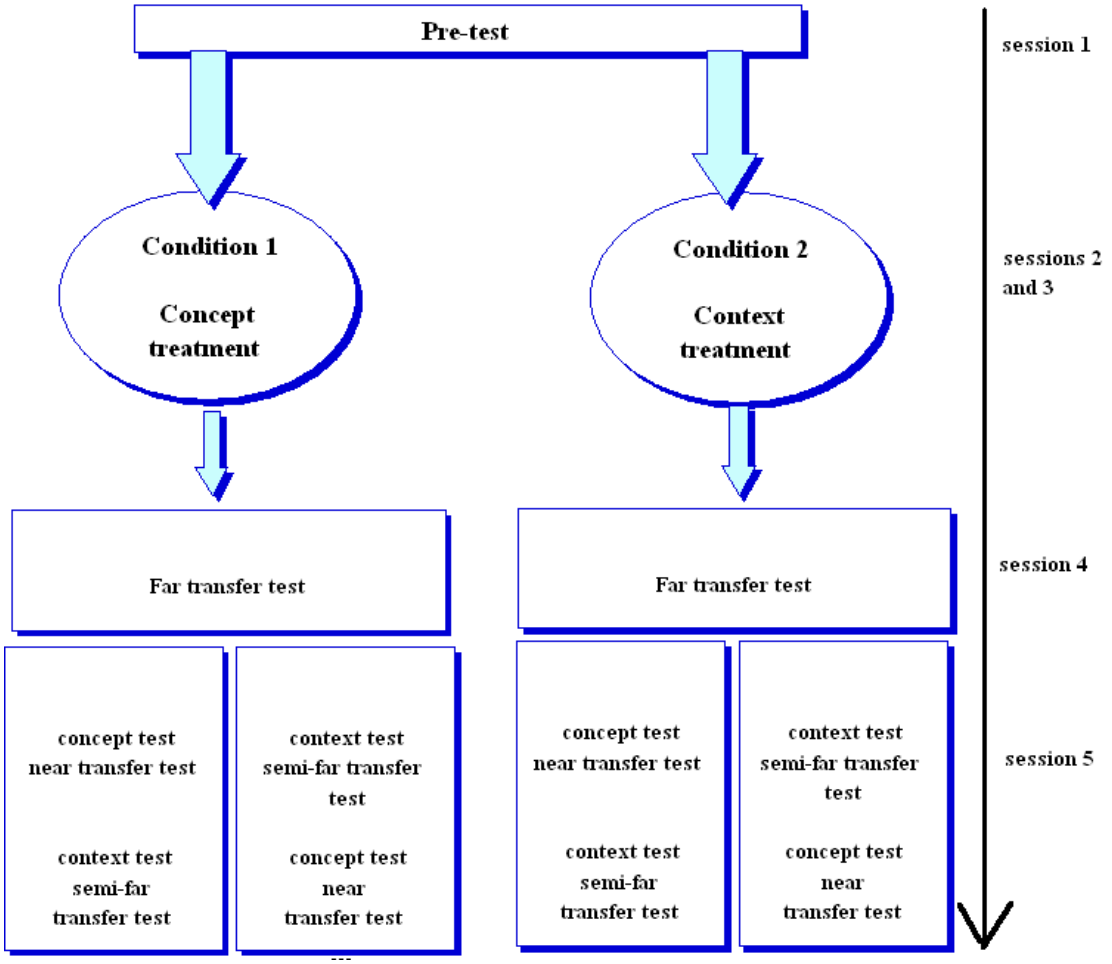
- a context post-test to measure the ability to make connections between concept and context (a near-transfer test for the context condition and a semi-far transfer test for the concept condition). This test consists of two open questions related to practical problems. One of the problems was 'abolition of the fiscal interest deduction for mortgages' that started from the practical context; the other was 'the economic climate in 2003', starting from the concept. Students were stimulated to make connections between context and concept and *vice versa*. Two independent researchers scored independently the connections between C and T in the students' texts. The minimum score was 0 and there was no maximum. The inter-judge reliability was good (Cohen's kappa 0.85).

Experimental procedure

Students were randomly assigned to one of the two conditions. The experiment was conducted within the school schedule of two lessons, distributed over the week. The tasks were self-directed. Explanations by the teacher, which could have been of influence on the performances of the students, were not necessary. The tasks and the tests were done via computer while students received all the information on paper. The instructions were presented by the same test leader in all groups.

Figure 6 represents the experimental procedure. The time schedule is indicated on the right. Each session lasted 50 minutes. In session 5 the concept and context tests were counter balanced. During the experiment sound recording took place by some of the pairs in both conditions. The researcher kept a logbook.

Figure 6: Course of the experiment for students in the concept condition and the context condition



Hypotheses

We expected that students in the concept condition would perform better on the concept post-test (a near transfer test for them) than students from the context condition (a semi-far test for them). We expected that students in the context condition would perform better in connecting concept and context (context post-test, a near transfer test for them) than students in the concept condition (a semi-far test for them). Furthermore, we expected that the concept condition would be a better starting point for solving far transfer tasks than the context condition. This expectation is based upon two relevant points from literature about meaningful learning (Alexander, 1997; Gelman and Greeno, 1989; Salomon and Perkins, 1989). First, a solid, coherently organised conceptual framework is a condition for transfer. Second, as far as we know, there is no empirical evidence to expect that the ability to connect concepts and contexts alone, without a strong conceptual network, is a good basis for far transfer.

5. Results

At the start of the experiment there appeared to be no significant difference between the conditions in the school marks for economics (t = –1.09; df = 1.34.5; p = 0.28) as in the results of the pre-test (t = –0.22; df = 138.1; p = 0.83).

Table 1 shows the students' scores in the concept and the context condition.

Table 1: Scores for the concept and context condition: means and standard deviations (between the brackets)

	Concept condition (N= 69) Mean	(SD)	**Context condition** (N=70) Mean	(SD)	**Max. score**	**Highest score**
Pre-test concept	40.16	(6.46)	40.47	(7.33)	72	
Post-test concept near/semi-far transfer (is equal to pre-test)	41.93	(8.20)	42.42	(7.95)	72	
Post-test context near/semi-far transfer	4.34	(3.46)	4.18	(3.40)	*	12
Post-test concept far transfer	46.60	(22.55)	49.03	(25.44)	*	85

* No maximum score

The mean scores on the pre-test show that the knowledge students possess of the concepts of money circulation, the state of the economy with related concepts and the relationships between these concepts, is low.

We expected that students in the concept condition should perform better on the near transfer concept test than students in the context condition, but they did not, as was proved by the results of the covariance analysis (controlled by the pre-test score) (F(3,132) = 0.059; p = 0.94). We expected also that the context instruction should result in making more connections between context and concept than the concept condition. However, the differences were not significant (F(3,129) = 0.356; p = 0.70). Concerning the far transfer, it was not proved significant that the concept condition resulted in more conceptual knowledge (F(3,136) = 0.213; p = 0.80) and more connections, (F(3,135) = 0.168; p = 85), compared to the context condition. From Table 1 we see that the scores on the context part of the far-transfer test, compared to the scores on the concept test, are low in both conditions. We expected this for the concept condition, but we did not expect that there would be no difference with the context condition. The instruction for the context condition was directed to making the connections.

Because the pre-test concept was the same as the near/semi-far transfer test concept, it was possible to determine whether students made progress on conceptual knowledge. From a t-test for a paired-sampled test we learnt that on the concept post-test, the mean scores were significantly better than on the pre-test (M/SD pre-test = 40.41 (6.01), M/SD post-test = 42.23 (6.96) (t = –2.663; df = 132; p = 0.009). The students improved their knowledge.

6. Discussion

In this study we compared the effects of two roads of instruction to reach near transfer, semi-far transfer and far transfer. The ability to transfer is an important goal to reach in economic education. To be able to achieve transfer students must have a deep understanding of economic concepts *and* they must be able to make connections between contexts and economic concepts. We designed two instructions: for the concept condition, the concept road emphasising the strengthening of economic concepts and for the context condition, the context road emphasising connections between concepts and contexts. We tested on near transfer, semi-far transfer and far transfer.

We could not confirm that the concept condition performed better on the near concept test and that the context condition performed better on the near context test. Also, we could not confirm that the concept condition, due to the better organised knowledge base, performed better on the far transfer test. But the whole group of students significantly improved in conceptual knowledge. Students had more knowledge of economic concepts after the two lessons. We expected that students in the concept condition, because of better concept knowledge, should be able to perform better on the far transfer test and that they should score better on the context part of the test. This was not the case. Most likely, the ability to make connections between context and concept is a necessary skill for far transfer and this has to be developed besides the conceptual knowledge. But in this case we should expect that the context condition that improved in conceptual knowledge as much as the concept condition should perform better on the context part of the far transfer test. Students have had an instruction aimed at making connections between concepts and contexts. Students in the context condition have not learnt more than students in the concept condition. The conceptual learning was the same in both conditions, but the ability to make connections was difficult in both conditions. A reason for the inability of far transfer in both conditions can be that, despite the extra training, the conceptual knowledge in both conditions is still not sufficient to reach far transfer. How can we explain this? A possible answer is that two lessons of 50 minutes each were too short to reach the aims. Possibly, the result can also be explained by the fact that the students followed the instructions only partly. They were asked to write down questions they engaged in performing the task, but they did not know what to ask and did not use the information they had received. Perhaps they needed more feedback than we provided in this experiment. Bransford and Schwarz (1999) underline the importance of new information and feedback. Working with context tasks can stimulate students to ask for feedback and information.

In spite of the results, this study has given us more insight, especially with regards to the following points:

- Conceptual knowledge alone is not sufficient for transfer.
- The ability to make connections between practical situations and concepts is not being learnt easily by students. In any case, it is more time-consuming than two lessons.
- In context-directed economics education students learn as many concepts as in concept-directed economics education. Also, students learn concepts *while* working on contexts.
- Student motivation depends on the type of task. Audio tapes of pairs and logbook notations showed that students were very motivated for the lessons in this experiment. To the astonishment of their teachers, in both conditions the students worked together very intensively, were serious and performed all the tasks with pleasure. The computer program and the task to construct a concept map played important roles.

The results are raising questions too. One question, for example, is: 'Is it better to start right from the beginning with contexts in economic lessons?' The approach of first teaching concepts and then in the next lessons teaching how to apply them in contexts is perhaps not the most effective way. Perhaps it is better to start with students working on a context and then go deeper into using concepts. We were supported in this by Alexander's model of domain learning (2005) mentioned earlier. It is also possible that for some students, the concept road is effective – for example, for students already belonging to the *competence* stage – while for others the context road is more effective – for example, for students belonging to the *acclimation* stage.

One important point to mention is that we have to ask ourselves if far transfer is possible in secondary education. Studies, where transfer has been measured, showed that often transfer was not reached (Detterman and Sternberg, 1993). Bransford and Schwarz (1999) attribute this to the fact that in school expert behaviour of students is often required. However, as they call it, direct application (DA) of knowledge is too difficult for students. They place more emphasis on preparation for future learning (PFL) by setting the goals on how to tackle the problem, how to question the problem, etc. This does not go as far as DA; it is the first step in reaching transfer.

7. Recommendations for future research

It is important that students can use their economic knowledge in daily life even when contexts are different from the contexts in which they have learnt the knowledge. Further research has to be done to find learning activities that lead to different forms of transfer, especially near and semi-far transfer. More research has to be done with respect to the use of the context road in economic education. In general, too little attention has been given to the linking of context to concepts, for which the context road can be a good method. In addition, it is motivating for students and it also helps them to develop conceptual knowledge. Perhaps alternating a focus on contexts with a focus on concepts is most effective. Future research could study such alternation or combination. Furthermore, more research needs to be done to determine which kinds of contexts are effective in the learning process and at which moments.

References

Alexander, P. A. (1997). 'Knowledge-seeking and self-schema: A case for the motivational dimensions of exposition', *Educational Psychologist,* Vol. 32 (Special issue), pp. 83-94.

Alexander, P. A. (2003). 'The Development of Expertise: The Journey From Acclimation to Proficiency', *Educational Researcher,* Vol. 32(8), pp. 10-14.

Alexander, P. A. (2005). 'Teaching towards expertise', *BJEP Monograph Series, Pedagogy – Learning for teaching,* Vol. 2(3), pp. 29-45.

Alexander, P. A. (2006). *Psychology of learning and instruction,* New Jersey: Pearson Education.

Anderson, L. W. (2002). 'Curriculum Alignment', *Theory into Practice,* Vol. 41(4), pp. 255-260.

Becker, W. E. (2004). 'Good-by old, hello new in teaching economics' [Electronic Version]. Available at SSRN: http://ssrn.com/abstract=601501.

Beyerbach, B. A. and Smith, J. M. (1990). 'Using a computerized concept mapping program to assess pre-service teachers' thinking about effective teaching', Special issue: perspectives on concept mapping, *Journal of research in science teaching*, Vol. 27(10), pp. 961-71.

Bloom, B. S. (1968). 'Learning for Mastery', *Evaluation Comment,* Vol. 1(2), pp. 1-5.

Bransford, J. D., Brown, A. L. and Cocking, R. R. (1999). *How People Learn,* Washington, DC: National Academy Press.

Bransford, J. D., and Schwartz, D. L. (1999). 'Rethinking transfer: a simple proposal with multiple implications', in (A. Iran-Nejad- and P. D. Pearson, eds), *Review of research in education,* Vol. 24 Chapter 3, pp. 61-100, Washington DC: American Educational Research Association.

Brown, A. L., Collins, A. and Duguid, P. (1989). 'Situated cognition and the culture of learning', *Educational Researcher,* Vol. 18, 32-42.

Chinn, C. A., and Brewer, W. F. (1993). 'The Role of Anomalous Data in Knowledge Acquisition: A Theoretical Framework and Implications for Science Instruction', *Review of Educational Research,* Vol. 63(1), pp. 1-49.

De Vellis, R. F. (1991). *Scale Development Theory and Applications,* London: Sage Publications.

Detterman, D. L. and Sternberg, R. J. (1993). *Transfer on trial,* Norwood NJ: Ablex.

Engeström, Y., Miettinen, R. and Punamäki, R.L. (1999). *Perspectives activity theory,* Cambridge: Cambridge University Press.

Gelman, R. and Greeno, J. G. (1989). 'On the nature of competence: Principles for understanding in a domain', in (L.B. Resnick, ed), *Knowing, learning and instruction: Essays in honor of Robert Glaser* (pp. 125-186), Hilsdale NJ: Laurence Erlbaum Associates.

Gorodetsky, M., Keiny, S., Barak, J. and Weiss, T. (2003). 'Contextual pedagogy: Teachers' journey beyond interdisciplinarity', *Teachers and teaching: theory and practice,* Vol. 9(1), pp. 21 -33.

Hansen, W. L., Salemi, K. M. and Siegfried, J. J. (2002). 'Use it or lose it: teaching Literacy in the Economics Principles Course', *American Economic Review,* Vol. 92(May), pp. 463-72.

Krathwohl, D. A. (2002). 'A revision of Bloom's Taxonomy: An Overview', *Theory into practice,* Vol. 41(4), pp. 212-18.

Marini, A. and Genereux, X. (1995). 'The challenge of teaching for transfer', in (A. McKeough, J. Lupart and A. Marini, eds), *Teaching for transfer: Fostering generalization in learning,* Mahwah, NJ: Lawrence Erlbaum.

Mayer, R. E. (2002). 'Rote versus Meaningful Learning', *Theory into Practice,* Vol. 41(4), pp. 226-32.

Mayer, R. E. (2004). 'Teaching subject matter', *Annu.Rev.Psychology,* Vol. 55, pp. 715-44.

Novak, J. D. (1990). 'Concept mapping: a useful tool for science education', *Journal of Research in Science Teaching,* Vol. 27(10), pp. 937-49.

Novak, J. D. (2002). 'Meaningful learning: The Essential Factor for Conceptual Change in Limited or Inappropriate Propositional Hierarchies Leading to Empowerment of Learners', in (G. J.Kelly and R.E. Mayer, eds), *Learning,* pp. 548-71, Wiley Periodicals, Inc.

O'Donell, A., Dansereau, D. F. and Hall, H. (2002). 'Knowledge maps as scaffolds for cognitive processing', *Educational Psychology Review,* Vol. 14(1), pp. 71-86.

Perkins, D. N. (1992). 'Transfer of learning', in *International Encyclopedia of education, second edition,* Oxford: Pergamon Press.

Pijls, M., Dekker, R. and Van Hout-Wolters, B. (2003). 'Mathematical level raising through collaborative investigations with the computer', *International Journal of Computers for Mathematical Learning,* Vol. 8, 191-213.

Roth, W. and Roychoudhury, A. (1992). 'The social construction of scientific concepts or the concept map as a conscription device and tool for social thinking in high school science', *Science Education,* Vol. 76(5), pp. 531-57.

Salomon, G. and Perkins, D. N. (1989). 'Rocky Roads to Transfer: Rethinking Mechanisms of a Neglected Phenomenon', *Educational Psychologist,* Vol. 24(2), pp. 113-42.

Stark, R., Mandl, H., Gruber, H. and Renkl, A. (1999). 'Instructional means to overcome transfer problems in the domain of economics: empirical studies', *International Journal of Educational Research,* Vol. 31(7), pp. 591-609.

Sternberg, R. J. (2003). 'What is an "expert student?"', *Educational Researcher,* Vol. 32(8), pp. 5-9.

Vosnadiou, S. (1994). 'Towards a revised cognitive psychology for new advances in learning and instruction', *Learning and instruction,* Vol. 22, pp. 45-69.

Whitehead, A. N. (1957). *The aims of education and other essays,* New York: Macmillan (Original published in 1929).

Appendix 1: A concept treatment assignment

Today, you are going to make a drawing – a diagram – that shows how the circular flow of income and expenditure is connected to the state of the economy/market conditions.

You already know all the concepts that belong to this topic from the classes of the preceding years. But it is useful to refresh your memory and to find out what you remember about this topic.

1. Do you know what all those concepts mean?
2. Do you know what the connections between those concepts are?
3. Can you account for and explain those connections?

If you are having trouble with this, it will be difficult for you to make sense of what is being said and written about the development of the economy right now. And that is quite a lot!

By means of this assignment you will discover what you have understood from the classes about this topic and on which points you still have work.

In these sessions we will work towards a poster-exhibition. We will use the materials you are going to make in the sessions for this. Today we are going to make the first set of materials.

What are you supposed to do?

You are going to make a diagram, using a list of concepts, a sheet of A3 paper, Post-it notes and a marker.

How are you going to go about this?

1. You have 10 minutes to do this individually.
2. In front of you there is a sheet of A3 paper and a pad of Post-it notes. On page 2 you will find the list of concepts.
3. Write down each concept on one Post-it note.
4. Paste the Post-it notes with the concepts on them on the sheet of paper in such a way that the concepts that are closely connected to each other end up close to each other on the sheet. Put the other concepts further away.
5. Draw lines or arrows between the concepts that you think are connected to each other.
6. Write down near the lines or arrows what the nature of the relationship between the concepts is. Use the markers for this.
7. You do not have to use all of the concepts. If you decide not to use certain concepts, please explain your reasons for omitting them. Put these concepts at the bottom of the sheet and write down why you haven't used them.
8. You are allowed to add concepts. You have empty post-it notes at your disposal. Please explain and write down why you have decided to use additional concepts.

Appendix 2: A context treatment assignment

Today, you are going to work on a current economic problem. By making the assignments you will complement your knowledge and you will learn to use your economic knowledge. And in practice it is necessary that you are able to do this. How else will you be able to join in on conversations about economic developments? In these sessions we will work towards a poster-exhibition. We will use the materials you are going to make in the sessions for this. Today we are going to make the first set of materials.

What is the current economic problem you are going to work on?

In the period running from May 2002 until October 2002 the Dutch government consisted of a coalition of the political parties CDA, VVD and LPF. This was rather extraordinary, since the LPF was a new party. This party was bent on changing numerous aspects of (political) life in the Netherlands.

The minister of Economic Affairs of the LPF, minister Heinsbroek, wanted to have a number of measures implemented as soon as possible. For example, in the spirit of the late LPF-leader Pim Fortuyn, who was assassinated shortly before the 2002 Dutch parliamentary elections, Heinsbroek wanted to give 'the Kwartje van Kok' back to the people. This term refers to the raising of the price of petrol by means of a supplementary duty of 25 cents (in Dutch guilders) per litre. This was carried out by the Dutch government under prime-minister Kok in 1991 to discourage motorists from driving. The LPF not only wanted to abolish this supplementary duty, but they also wanted to give this money back to the people retroactively. This would mean that Dutch families receive a sum of money. At the same time this would cost the government a great deal of money.

Of course, minister Heinsbroek had read in the 'Macroeconomic Explorations' of the Dutch Central Planning Bureau that the economic growth had declined in 2002. It was a matter of underspending and he was very much aware of that. He also knew that in the following years the economic growth would remain limited or would potentially even decline further.

At that time many people were wondering: These plans of minister Heinsbroek, are they full of hot air, doomed to deflate sooner or later, or are these measures useful in the current economic depression?

You have been trained in economics. You understand that this question cannot be answered with a simple yes or no. What are the consequences of the repayment of the money for the Dutch families? How would that influence the economic situation of that time? And what are the consequences of the repayment of the money for the Dutch government? How would that influence the economic situation of that time? And how do these two factors work together? Will the repayment of 'the Kwartje van Kok' help to get out of the economic depression or will it only make things worse?

What is your answer when you are presented with this problem?

What are you supposed to do?

During a dinner-party with your friends the question explained above suddenly comes up!

These plans of minister Heinsbroek, are they full of hot air, doomed to deflate sooner or later, or are these measures useful in the current economic depression/slump?

And everyone is looking at you; after all, you are taking economics! You do not have much time to think about it. What are you going to answer?

Write down in the box below the answer you would have given in the situation just described.

Appendix 3: Five economic circumstances are indicating poor performance in the economy

The economic editor of the journal has written an article about five economic circumstances that together are indicating poor performance in the economy.

Each part of the article starts with a picture with a title. The five titles are:

1. Exports are going down…..
2. …*and* consumer confidence and producer confidence are collapsing…
3. …*so* economic growth is decreasing….
4. …*as a result of this* unemployment is increasing…
5. …*and* the budget deficit is growing.

With every image the editor had written a paragraph to explain to the newspaper readers *what influence the mentioned circumstances had for their daily lives*, why the circumstances developed and how they influence subsequent events. The words *and… so…..as a result…* show the causality.

But what happened? The editor has lost his writings and he has left for holiday to an unknown destination. What to do?

You will understand. The editor is convinced about your capacity to complete the article by writing the lacking parts. With the information you will find underneath, you will succeed!

Get to work as follows:

Write the texts, that go together with the images, for readers that not have much knowledge of the economy.

So write in your own words:

- what the image is showing
- why the image is in the position it is
- why the circumstance is proving that poor performance is expected and what consequences this has for the daily lives of the readers

Each text must have a minimum length of 100 words.

Good luck!

Appendix 4: Procedure for scoring connections between conceptual knowledge and practical context

Three steps were taken.

1. Division into units.

The answers of the students were broken up into the following meaningful units (scoring units).

Sentences
Example: *For most people the purchase of a house is an enormous expenditure they do not make overnight.*

Parts of sentences, if such a part contains a new point.
Example: *If they do not want to do this, houses will continue to be for sale and this will be the beginning of a depression.*

Conjoined sentences – when a statement in one particular sentence is explained in the next sentence.
Example: *This will also have an effect on the market conditions. These will start to decline because potential buyers of houses will start to lose faith in the economy.*

2. Attribution of T- *and* C-*scores for units that concerned 'practical context' and 'economic concepts' respectively.*

Code *T* is attributed:

- When a scoring unit renders a description/aspect of a context. For example: *When this is no longer possible, house owners may end up in a tight spot because they are no longer able to pay off the mortgage.* Or when a statement is given that is in itself economic, for example: *If this is the case the market conditions will enter a period of recession,* but that lacks any further explanation. In such a unit a student delivers a statement (*show**). A statement is often followed by an explanation and/or justification (**explain and justify*) (Dekker and Elshout-Mohr, 1996). An explanation can be recognised by the presence of words such as: *so, due to this, because of, as a consequence of, therefore,* etc. When the explanation is provided in a context this is scored *with* a *T*, for example: *Due to this, the market conditions will also start to decline again, because the person in question has to pay more (will have to pay more) (for a house) (explain and justify).*
- When economic concepts are used in the explanation, a *C* is attributed. For example: *This is bad for the economy. Because people will start to limit their expenditures.*

3. Computation of the total score.

The number of transitions a student made between context and concept or *vice versa* is added up.

Example: Peter's answer to one item of the Context-test is presented in italic; each unit (defined as unit by the experimenter) is presented at a separate line.

T *consumer loses more money to mortgage*
C *consumer's purchasing power declines*
T *consumer moderates buying behaviour; buying is expensive*
C *invests less and saves more*
T *buys fewer products*

- ***C*** *producers' faith declines*
- ***C*** *production capacity is not fully exploited*

Peter makes 5 transitions (*T*-*C*, *C*-*T*, *T*-*C*, *C*-*T* and *T*-*C*) and obtains a score of 5.

Contact details

Lenie Kneppers
University of Amsterdam
Spinozastraat 55
1018 HJ Amsterdam
The Netherlands
Email: h.c.kneppers@uva.nl

Peer Effects on Undergraduate Business Student Performance

Salvador Contreras, Frank Badua and Mitchell Adrian

Abstract

We evaluate the endogenous peer influence of students in one U.S. public University College of Business. In particular, we measure the peer effect on student grades. This study utilises an exclusion restriction approach similar to De Giorgi *et al.* (2010) to estimate the endogenous peer effect. Our results support the finding that a student's classroom performance has a significant effect on their peers. Overall, our results suggest a negative peer effect. However, we find that the direction and magnitude of the peer effect is sensitive to the student's own average ability and that of their peers.

JEL classification : A10, A13, A22, C70, Z13

1. Introduction

Research has shown that social interactions play an important role in influencing human behaviour. Studies on peer effects cover a wide spectrum of social contexts. For example, some analyse the role peers have on body weight (Trogdon, *et al.*, 2008), substance abuse (Kawaguchi, 2004) and academic performance (Sacerdote, 2001). Our paper focuses on the role peers have on one's academic performance. Social interactions that influence a student's ability to learn has long been of interest to researchers, policy makers, and parents. Researchers have found this question difficult to answer given the many environmental variables that can influence student performance. Even when reliable data are found there are identification issues that arise in evaluating peer effects. Manski (1993) outlines the problem of identification in the literature as the ability of researchers to separate peer effects from endogenous, exogenous or correlated effects.[1]

Recently, Lee (2007) employing a network interaction strategy shows that with groups of moderate size and sufficient size variation identification can be achieved. A number of recent papers have built on Lee's (2007) work (Bramoulle *et al.*, 2010; Boucher *et al.*, 2010) to disentangle endogenous effects from correlated effects. The use of group size variation has however been used in the literature with mixed results in determining peer effect significance (Ammermueller and Pischke, 2009; Hoxby, 2000). Recently, De Giorgi *et al.* (2010) used an exclusion restriction identification strategy that estimates indirect peer effects. The exclusion restriction approach makes use of the causal effect that one group has on another in the absence of direct contact. For example, group A may be influenced by group C, even though their members never meet, because members of A and C both meet and interact with members of group B, who serve as intermediaries of influence. Group C, whose members indirectly influence A, is known as the excluded group. De Giorgi *et al.* (2010) use a set of academic courses common to several peer groups to show that indirect peer effects influence a student's choice of college major.

[1] See Bramoulle *et al.* (2009) and Hanushek *et al.* (2003) for an expanded discussion of the issue.

In this paper, we evaluate data from a College of Business in a State University located in the southern United States. Our analysis consists of measuring the endogenous effect that the average performance of one group of peers has on another group of students. We employ an exclusion restriction identification approach (De Giorgi *et al.*, 2010). Our estimating strategy is to evaluate the effect that one group of students has on another but where there is no shared direct interaction. In our data we identify three groups of students, groups A, B and C. One group of students (group A) has taken a course with another group of students (group B). A third group of students (group C) has taken a course with members of group B. However, group C students have never taken courses with group A. As defined earlier, group C is also referred to as the excluded group. The estimating strategy is to measure the effect that group C has on group A through group C's direct interaction with group B. We find that a group of students' grade point average (GPA) is explained by the excluded group's GPA. Our results show that the endogenous peer effects are significant and negative. The construction of our model as well as the significance of the endogenous peer effect is consistent with De Giorgi *et al.* (2010). However, the direction of the effect found is contrary to some peer effects studies (De Giorgi *et al.*, 2010; Hanushek *et al.* 2003). Nonetheless, Zimmerman (2003), Sacerdote (2001), and Henderson *et al.* (1978) show that the direction of the peer effect may be a function of the interaction between one's own ability and that of peers.

Our results show that students' ability plays an important role in explaining the magnitude and direction of the peer effect. We find that the excluded group (group C) students who are of below average ability indirectly hurt the explained group's (group A's) average GPA. However, excluded group (group C) students of above average ability pulled up the GPA of students (group A's) of above average ability. It is difficult to derive strong policy implications from these findings since in our dataset course selection is not random. However, our results suggest that diversity in a group's ability works against the academic achievement of high ability students (Kim *et al.*, 2008).

The paper is organised as follows. We first describe the existing literature, then introduce our data, outline our identification strategy and present our estimating strategy and our baseline results. We then go on to analyse the role of a group's ability and its effect on the magnitude and direction of the peer effect. The final section offers some concluding remarks.

2. Existing literature

The literature on student performance is diverse. One branch of this literature has evaluated the effect of class size on student achievement. The general hypothesis is that lower student enrolment allows for greater access to teachers. Most prior literature has shown that small class sizes positively affect student performance (Krueger, 1999; Kokkelenberg *et al.*, 2008). However, there are some studies that find no effect (for example Hoxby, 2000). Another branch of research has focused on the effects of peer performance on a student's performance (Boucher *et al.*, 2010; Black *et al.*, 2010, Ammermueller and Pischke, 2009; Kang, 2007; Hanushek *et al.*, 2003; Henderson *et al.*, 1978), or evaluates the effects of peer abilities on student achievement (Kim *et al.*, 2008; Zimmerman, 2003; Sacerdote, 2001). This paper belongs to the latter group.

Work by Kim *et al.* (2008) comparing two alternative school systems in South Korea shows that when students with similar academic abilities are assigned to the same high school (sorting) this improves the average performance of top students as compared to school systems where students are not sorted. Furthermore, they find that sorting produces a significant peer effect on top students. However, they also find that average and low ability students derive no significant benefit from sorting.

In a different context, Sacerdote (2001) using data from Dartmouth College shows that random assignment of college roommates has a significant effect on student outcomes. This present paper is closer in spirit to that of Sacerdote (2001). One difference is that we have a smaller sample from a

commuter public regional university that is less selective than Dartmouth. As in Sacerdote (2001), we use GPA as our dependent variable and use ACT (American College Testing) scores and high school GPAs as measures of innate ability. However, our findings are somewhat different. The reasons are most likely due to differences between the two populations. Our sample is more heterogeneous in respect to innate abilities, social-economic background and quality of high school. More importantly, our sample is much more restrictive. We captured students in one College of Business who have successfully completed all degree requirements. To some extent, our estimates probably understate peer effects as we were unable to measure the effect that unsuccessful peers may have on student achievement.

As outlined in Boucher *et al.* (2010) and Ammermueller and Pischke (2009), an important consideration in the literature is distinguishing between various factors of student performance. The various causes may be exogenous (influences of peers' characteristics), endogenous (resulting from interactions within the peer group) or from common influences on the peer group (correlated effects) (Manski, 1993). The literature uses various strategies to isolate these effects. The research is generally categorised by natural experiments (Boozer and Cacciola, 2001; Sacerdote, 2001; Zimmerman, 2003; Kang, 2007), quasi-experimental (Carrell *et al,.* 2009; Hoxby; 2000), fixed effects (Vigdor and Nechyba, 2004; Hanushek *et al,.* 2003) or network studies (Bramoulle *et al*., 2010; Boucher *et al.*, 2010; De Giorgi *et al.*, 2010; Lee, 2007). This paper belongs to the latter category.

3. Data description

We collected data from a public university's College of Business located in the southern United States. All students in the sample were declared business majors who completed all coursework required for graduation. The data were originally collected to evaluate and predict student performance on the Business Major Field Test (MFT) exam administered by the Educational Testing Services (ETS). The data consist of MFT scores, undergraduate GPA, high school GPA, ACT scores and demographic information. The MFT is taken in the last semester of coursework leading to a Bachelors of Science in one of five business programmes. The MFT is embedded in the capstone course, which is offered every semester and is restricted to graduating seniors. The data consist of six semesters with MFT score information. Each student's MFT score is accompanied with two years of detailed coursework information.

Table 1A: Descriptive statistics

	Mean	SD	Min	Max
Gpa	3.01	0.57	1.3	4
act_comp	20.79	3.45	12	28
hs_gpa	3.26	0.53	1.8	4
White	0.82	0.38	0	1
Black	0.13	0.34	0	1
Female	0.52	0.50	0	1
Age	23.22	3.03	20.6	44.6
Accounting	0.32	0.47	0	1
Finance	0.11	0.31	0	1
Gen Business	0.11	0.32	0	1
Management	0.24	0.43	0	1
Marketing	0.23	0.42	0	1

We have detailed grade information on the last two years of study leading to a bachelor's degree. Table 1A has descriptive statistics of selected variables in our dataset. Table 1B provides variable definitions for those variables reported in Table 1A. Students in their last two years of study had an average GPA of 3.0. One student completed his last two years of college work with a GPA of 1.3. The average ACT

composite score is 20.8. On average, students were 23 years of age at the time they completed their course requirements leading to a bachelor's degree in business. The oldest student in the sample is 45 years of age while the youngest is 21. About 95% of the students in the sample are 26 years old or younger. The sample is composed of 52% female and 13% African American students. Roughly, 32% of all students in our sample were accounting majors, followed by management (24%). The least popular majors were finance and general business with 11% each.

Table 1B: Variable definitions

Variable name	Description
Gpa	Average GPA of student i. Last 2 years of college courses.
act_comp	ACT composite score of student i.
hs_gpa	High school GPA of student i.
White	White race dummy.
Black	African American race dummy.
Female	Female gender dummy.
Age	Age at time of MFT exam.
Accounting	Accounting major dummy.
Finance	Finance major dummy.
Gen Business	General Business major dummy.
Management	Management major dummy.
Marketing	Marketing major dummy.

4. Peer groups and identification strategy

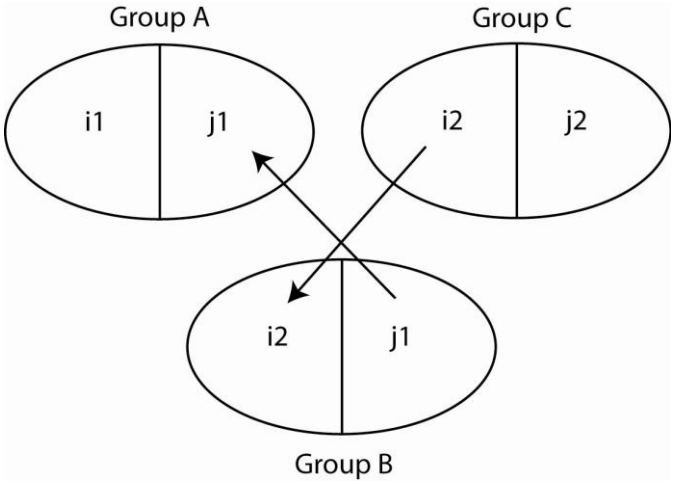
Our dataset identifies two courses that all business students took for credit. The capstone course was taken by all business students in their last semester. The second course is a junior level Operations Management (OM) course. On average, students took the OM course one semester before their final semester. Some 29% took the OM course in their final semester. No student in our database took the OM course more than three semesters away from their last. Our dataset contains 22 OM and 15 capstone sections.

We propose an identification strategy similar to that outlined in De Giorgi *et al.* (2010). We group students based on class pairings for the OM and capstone course only. We had 322 students in our dataset that took these two courses for credit. Each student *z* chooses a course *i* and *j* to fulfil degree requirements. Let course *i* represent OM and *j* capstone courses. Each student *z* is then grouped in a pair of courses *ij*. There were 122 unique pairs of *ij* courses with an average group size of 42.5 students.

We then identified groups that had direct connections to another and all indirect groups. Because the OM course is always taken before the capstone course we paired the peer and the indirect group by course *j*. Each group had a set of peer groups and their respective group of indirect groups. Figure 1 illustrates this relationship. We generated a total of 5,069 unique three-way group pairings. Each pairing had information on the average GPA, percentage of African American, percentage of female, average high school GPA, average ACT composite score and average age of students of the two-course pair.

In Figure 1 we can observe that group A and group B share a common course. Similarly, group B and group C share a common course. However, group A and the excluded group (group C) do not share a common course. The identification strategy outlined in De Giorgi *et al.* (2010) suggests that if a peer effect exists it should be observed indirectly. That is, if group C influences group A it must be through group C's influence on group B. This is the exclusion restriction approach taken in our empirical analysis.

Figure 1: Indirect effect of excluded group on group A



The group pairings allows us to control for non-peer effects such as course date, time, instructor, etc. However, this setup may overstate the peer effect. First, we are not able to observe non-classroom interactions between group A and group C students. For example, marketing students may be members of the American Marketing Association. Secondly, some group A students may have taken other courses with group C students that we ignored in the construction of three-way pairings. Thirdly, assignment to pair courses are quite likely endogenous. Putting the limitations of our dataset aside, we believe that our overlapping interaction group and identification strategy provide a compelling case for reducing the effect of the reflective problem as outlined in the literature.

5. Estimating strategy

The strategy is as follows. There are two courses that all students take. Course *i* (Operations Management) that is taken before course *j* (capstone course). Both courses are populated by three overlapping groups of students. Group A takes course *i*1 and *j*1, group B takes course *i*2 and *j*1, and group C takes course *i*2 and *j*2. Group A and group B share the same capstone course *j*. While, group B and group C share the same Operations Management course *i*. However, no interaction takes place between group A and group C. If a peer effect exists, it should be identified indirectly. That is, if group C has an influence on group A it must be through the effect group C had on group B.

We propose the following two-stage estimating equations:

(1) $gpa_{B,j} = \beta_1 + \beta_2 gpa_{C,j} + \beta_2 X_{C,j} + u_j$

(2) $gpa_{A,i} = \alpha_1 + \alpha_2 gpa_{B,j} + \alpha_2 X_{A,i} + u_{i,j}$

In the first stage, we estimate the average GPA of group B instrumented by the average GPA of group C and group C specific characteristics ($X_{C,j}$). In the second stage, we estimate the average GPA of group A with the fitted GPA of group B and group A's specific controls ($X_{A,i}$). If a peer effect exits, then α_2 should be significantly different from zero.

6. Results

From equations (1) and (2) we produce the regression results reported in Table 2. We show in Column (1) the OLS specification for comparison purposes. The OLS results show that the average GPA of peers has a significant positive effect on group A. Columns (2) and (3) are our 2SLS specifications. Column (3) uses group C's GPA as an instrument for group B's GPA. Column (3) includes the full set of instruments.

Recall that the 2SLS specification tests whether group C influences group A through group C's effect on group B. Columns (2) and (3) show that group C's influence on group B lead to a negative effect on group A's GPA. We discuss the reason for the negative sign in the next section. However, it is important to point out that the effect is significant. That is, our exclusion restriction method supports the existence of a peer effect. In column (3) specification, group B's GPA is estimated (equation (1)) by group C's GPA and group C's specific characteristics. This lowers the magnitude of the effect only slightly. To evaluate the magnitude of peer effect we standardise our estimate as in Ammermueller and Pishke (2009). We multiply the first stage fitted variance of the peer group by the peer coefficient and divide it by the variance of group A's GPA. This procures an effect of –0.09. This suggest that a 1 standard deviation increase in the fitted peer groups GPA lowers group A's GPA by 0.09 standard deviations, where the mean GPA of group A is 2.89 and its variance is 0.02.

Table 2: Peer effect on group performance

Dependent variable:	OLS	2SLS	2SLS
Group A's average GPA	(1)	(2)	(3)
Peer effect:	0.292***	–0.218***	–0.199***
Group B's GPA	(0.014)	(0.031)	(0.027)
Group A's characteristics			
Percentage African American	–0.307***	–0.289***	–0.290***
	(0.033)	(0.037)	(0.037)
Percentage female	0.439***	0.683***	0.674***
	(0.023)	(0.029)	(0.028)
Average high school GPA	–0.177***	–0.471***	–0.461***
	(0.028)	(0.035)	(0.033)
Average age	0.020***	0.030***	0.030***
	(0.003)	(0.003)	(0.003)
Average ACT composite score	0.025***	0.042***	0.041***
	(0.004)	(0.005)	(0.005)
Constant	1.407***	3.130***	3.066***
	(0.143)	(0.185)	(0.176)
Observations	5069	5069	5069
R-sq	0.269	–	–
Shea partial R^2	–	0.254	0.336
First stage F-test	–	1646.8	425.11

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (standard errors in parentheses).

Group A's specific controls produce coefficients that are consistent with existing literature. Groups that are associated with a larger share of African Americans tend to achieve a lower GPA. Groups composed of a larger share of female students achieve higher GPAs. Here we find that higher group high school GPAs lead to a lower average GPA. However, this is offset by the average composite ACT score of the group that has a positive effect on average GPA. Lastly, groups composed of older students on average will have higher GPAs.

7. Excluded group ability

In this section we test the robustness of our results by the type of excluded group (group C's) ability level. We used the specification of Table 2 column (3) as our comparison estimation results. We take the average ability of the excluded group (group C) and break it into four categories (bottom 25th percentile, bottom half, top half, and top 75th percentile). We produce estimated results of these categories in Table 3. Column (1) shows the effect of the excluded group students (group C) from the bottom 25th percentile of the ability spectrum, on group A students. Column (2) shows the effects of group C students from the bottom 50th percentile of ability, on group A students. Both columns (1) and (2) show that low ability excluded group students (group C) have a negative effect on group A's GPA. The effect, while negative, is not significant for high ability excluded group (group C) students (columns (3) and (4)).

The results reported in Table 3 suggest that the negative effect found on our baseline estimation is probably driven by low ability excluded group (group C) students. Columns (1) and (2) peer effect is also a larger negative number than our baseline model. This suggests that low ability excluded group (group C) students drive most of the negative peer effect on group A students.

Next, we further partition our data to evaluate the effect of the excluded group's (group C) ability by groups A's ability level. Again, we take the average ACT composite score of group A as the measure of ability. Table 4 columns (1) and (2) show the bottom half ability of the excluded group (group C) on the bottom half (column (1)) and top half ability group A students (column (2)). Columns (1) and (2) clearly show that low ability excluded group (group C) students pull down both low ability and high ability group A students. However, the magnitude of the negative effect is larger on higher ability group A students.

Table 3: Peer effects on group performance – by excluded group ability levels

	Excluded group ability			
	< 25th Percentile	Bottom half	Top half	> 75th percentile
Dependent variable:	2SLS	2SLS	2SLS	2SLS
Group A's average GPA	(1)	(2)	(3)	(4)
Peer effect:	–0.210***	–0.215***	–0.020	–0.010
Group B's GPA	(0.044)	(0.029)	(0.043)	(0.060)
Group A's characteristics				
Percentage African	–0.077	–0.207***	–0.361***	–0.321***
American	(0.079)	(0.051)	(0.050)	(0.065)
Percentage female	0.751***	0.717***	0.525***	0.511***
	(0.048)	(0.034)	(0.044)	(0.061)
Average high school	–0.442***	–0.443***	–0.382***	–0.433***
GPA	(0.067)	(0.045)	(0.047)	(0.061)
Average age	0.015***	0.027***	0.029***	0.033***
	(0.006)	(0.004)	(0.004)	(0.005)
Average ACT	0.047***	0.043***	0.034***	0.031***
composite score	(0.010)	(0.007)	(0.007)	(0.008)
Constant	3.199***	3.044***	2.549***	2.645***
	(0.350)	(0.230)	(0.254)	(0.345)
Observations	1313	2845	2224	1301
Shea partial R^2	0.493	0.45	0.277	0.277
First stage F-test	210.66	386.74	141.29	82.12

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (Standard errors in parentheses).

Columns (3) and (4) consider the effect of the group C students from the top half of the ability spectrum on group A students from the bottom (column (3)) and top (column (4)) halves of the ability spectrum. Column (3) shows that top half ability excluded group (group C) students pull down bottom half group A students. This result is hard to reconcile. One possible explanation may be that high ability group C students discourage middle to low ability group B students who in turn discourage low ability group A students. Column (4) shows that high ability excluded group (group C) students indirectly influence high ability group A students in a positive way. This effect is of the same magnitude as that found in column (2), albeit of different signs.

Table 4: Peer effect on group performance – by excluded group and group A's ability levels

	Excluded group and group A's ability			
	Group C: bottom half ability		Group C: top half ability	
Dependent variable: Group A's average GPA	Group A: bottom half ability 2SLS (1)	Group A: top half ability 2SLS (2)	Group A: bottom half ability 2SLS (3)	Group A: top half ability 2SLS (4)
Peer effect:	–0.099***	–0.343***	–0.224***	0.334***
Group B's GPA	(0.032)	(0.051)	(0.062)	(0.045)
Group A's characteristics				
Percentage African	–0.466***	0.548***	–0.646***	0.198***
American	(0.072)	(0.090)	(0.078)	(0.074)
Percentage female	0.444***	1.058***	0.408***	0.451***
	(0.043)	(0.057)	(0.056)	(0.058)
Average high school	–0.517***	–0.543***	–0.646***	–0.272***
GPA	(0.059)	(0.071)	(0.075)	(0.056)
Average age	0.068***	0.044***	0.071***	0.050***
	(0.014)	(0.014)	(0.016)	(0.012)
Average ACT	–0.021***	0.086***	–0.016***	0.056***
composite score	(0.005)	(0.006)	(0.006)	(0.005)
Constant	3.841***	1.993***	4.504***	0.109
	(0.347)	(0.405)	(0.420)	(0.318)
Observations	1554	1291	1241	983
Shea partial R^2	0.552	0.391	0.283	0.386
First stage F-test	316.75	136.95	80.95	101.7

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (Standard errors in parentheses)

8. Conclusion

In this paper, we evaluated peer effects using an exclusion restriction identification approach similar to that employed by De Giorgi *et al.* (2010). Results of this research indicate that student influence on peers takes place in the absence of direct contact in a common course. Additionally, in this specific setting, our results suggest that the excluded group (group C) affects students indirectly in a negative way. The negative effect is primarily driven by the ability type of student groups. Low ability excluded group (group C) students had a negative indirect effect on other students. While, high ability excluded group (group C) students had a negative impact on low ability students. Finally, high ability excluded group (group C) students had a positive influence on other high ability students. Results suggest that endogenous peer effects are significant and negative for students with low academic aptitude relative to peers. However, we found no significant peer effect for high ability students (group A) among low

ability peers. Our results support existing literature that find that peer group performance has a significant effect on student academic achievement.

It is difficult to draw policy implications from these types of studies. In particular, work by Manski (1993) and more recently Bramoulle *et al.* (2009) have shown that parsing out the effects of groups on individuals is a challenging task. Nonetheless, the present study suggests that there is an intimidation factor present whenever low ability students are among high ability students. It suggests that classroom diversity may not be an appropriate tool to draw positive GPA externalities from social interactions. These results support a sorting strategy as found in Kim *et al.* (2008).

References

Ammermueller, A. and Pischke, J. (2009). 'Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study', *Journal of Labor Economics*, Vol. 27(3), pp. 315–48.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2010). 'Under Pressure? The Effect of Peers on Outcomes of Young Adults', IZA Discussion Paper 4946, Institute for the Study of Labor (IZA).

Boozer M. A. and Cacciola S. E. (2001). `Inside the 'Black Box' of Project STAR: Estimation of Peer Effects Using Experimental Data', Working Papers 832, Economic Growth Center, Yale University.

Boucher, V., Bramoullé, Y., Djebbari, H. and Fortin, B. (2010). 'Do Peers Affect Student Achievement? Evidence from Canada Using Group Size Variation', IZA Discussion Papers 4723, Institute for the Study of Labor (IZA).

Bramoullé, Y., Djebbari, H. and Fortin, B. (2009). 'Identification of peer effects through social networks', *Journal of Econometrics*, Vol. 150(1), pp. 41–55.

Carrell, S. E., Fullerton, R. L., and West, J. E. (2009). 'Does Your Cohort Matter? Measuring Peer Effects in College Achievement', *Journal of Labor Economics*, Vol. 27(3), pp. 439–64.

De Giorgi, G., Pellizzari, M. and Redaelli, S. (2010). 'Identification of Social Interactions through Partially Overlapping Peer Groups', *American Economic Journal: Applied Economics*, Vol. 2(2), pp. 241–75.

Hanushek E. A., Kain J. F., Markman J. M. and Rivkin S. G. (2003). 'Does peer ability affect student achievement?', *Journal of Applied Econometrics*, Vol. 18(5), pp. 527–44.

Henderson, V., Mieszkowski, P. and Sauvageau, Y. (1978). 'Peer group effects and educational production functions', *Journal of Public Economics*, Vol. 10(1), pp. 97–106.

Hoxby, C. M. (2000). 'The Effects Of Class Size On Student Achievement: New Evidence From Population Variation', *The Quarterly Journal of Economics*, Vol. 115(4), pp. 1239–85.

Kang, C. (2007), 'Academic interactions among classroom peers: a cross-country comparison using TIMSS', *Applied Economics*, Vol. 39(12), pp. 1531–44.

Kawaguchi, D. (2004). 'Peer effects on substance use among American teenagers', *Journal of Populations Economics*, Vol. 17, pp. 351–67.

Kim, T., Lee, J. and Lee Y. (2008). 'Mixing versus sorting in schooling: Evidence from the equalization policy in South Korea', *Economics of Education Review*, Vol. 27(6), pp. 697–711.

Kokkelenberg, E. C., Dillon, M. and Christy, S. M. (2008). 'The effects of class size on student grades at a public university', *Economics of Education Review*, Vol. 27(2), pp. 221–33.

Krueger, A. B. (1999). 'Experimental Estimates Of Education Production Functions', *The Quarterly Journal of Economics*, Vol. 114(2), pp. 497–532.

Lee, L. (2007). 'Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects', *Journal of Econometrics*, Vol. 140(2), pp. 333–74.

Manski, C. F. (1993). 'Identification of Endogenous Social Effects: The Reflection Problem', *Review of Economic Studies*, Vol. 60(3), pp. 531–42.

Sacerdote, B. (2001). 'Peer Effects With Random Assignment: Results For Dartmouth Roommates', *The Quarterly Journal of Economics*, Vol. 116(2), pp. 681–704.

Trogdon, J., Nonemaker, J. and Pais, J. (2008). 'Peer effects in adolescent overweight', *Journal of Health Economics*, Vol. 27, pp. 1388–99.

Vigdor, J. and Nechyba, T. (2004). 'Peer Effects in North Carolina Public Schools', Mimeo, Department of Economics, Duke University.

Zimmerman, D. J. (2003). 'Peer Effects in Academic Outcomes: Evidence from a Natural Experiment', *The Review of Economics and Statistics*, Vol. 85(1), pp. 9–23.

Contact details

Salvador Contreras
The University of Texas Pan American
Department of Economics and Finance
1201 W. University Drive
Edinburg, TX 78539, USA
Email: contrerass@utpa.edu

Frank Badua, Ph. D.
Lamar University
Department of Accounting and Business Law
P.O. Box 10069
Beaumont, TX, 77710, USA
Email: frank.badua@lamar.edu

Mitchell Adrian
Box 92232
McNeese State University
Lake Charles, LA 70609, USA
Office: 337-475-5535
Email: madrian@mcneese.edu

Using Online Courseware to Play a Simulation Illustrating the Concept of Moral Hazard in Health Care

Michael H. Kennedy

Abstract

This article discusses the use of a popular online instructional management system to teach moral hazard by having students make decisions about whether to seek health care under two simulated scenarios – with and without insurance.

JEL classification: A20, G22, I11

1. Introduction

Teaching economic principles to undergraduate and graduate students within the context of courses about health care systems has enormous value. Economic principles provide insights into how patients, providers and the institutions they represent behave. Insights gained by students in the classroom setting should help to better inform the decisions these students will make as health care leaders. Moral hazard is a particularly useful principle because of the implications about how those insured from loss due to illness or injury behave differently than those without insurance.

Classroom experiments, simulations, and games engage students in experiential learning about economic principles (Kagel and Roth, 1995; Friedman and Sunder, 1994), but they can be challenging to implement for students taught outside of the traditional classroom. Teaching students at a distance presents a number of challenges, and instructors can slip into the pitfall of focusing on the constraints rather than the opportunities (Larreamendy-Joerns and Leinhardt, 2006). The limitations of courseware designed to facilitate online instruction are amplified when used poorly (Foshay, 2002; Jonassen, 2002); however, as illustrated by the discussion of the moral hazard simulation that follows, the functions inherent in the software can be leveraged to produce teachable moments not easily replicated with face-to-face instruction. Although simply 'hanging content' on the Internet is a suboptimal practice that fails to take full advantage of the capabilities of internet courseware (Vrasidas, 2004), using an online course management system as a repository for lecture notes and other course content has migrated from distance education to become a common practice for all modes of course delivery (Brown, 2001).

Better practices involve using instructional software to experientially engage students in the consideration of alternatives and to promote active decision making. This article illustrates the use of a popular online instructional management system to have students make decisions about seeking health care under two scenarios, without and with insurance, during the play of a game designed by the author to illustrate the economic concept of moral hazard. In the context of health systems, moral hazard refers to the tendency for a person to seek health care in greater quantity and at greater cost with insurance than without. Intuitively, the concept makes sense, and this principle is invariably

illustrated in popular health economics and health systems texts (Feldstein, 2005; Folland *et al.*, 2010; Henderson, 2009; Morrisey, 2007), including the text used in one of the courses in which the online simulation is played to illustrate moral hazard (Getzen, 2010). Student play of the game represents an adaptation of experimental economics facilitated by the online courseware. Game play becomes a classroom experiment to determine whether students behave as predicted by the principle of moral hazard. Ortmann and Colandar (1997, p. 449) suggest that providing a classroom environment in which students experience economics has pedagogical advantages, and that 'experiments demonstrate to students subtleties that would otherwise be missed'.

The goals of this article are to:

1. describe how the online courseware was adapted to facilitate asynchronous, online play of the game by the students;
2. determine whether the results are reflective of behaviour influenced by moral hazard; and
3. determine the strategies used by the students while playing the game without and with insurance.

2. Playing the game

Set-up

The game used to illustrate moral hazard is played in two parts – Part 1 is entitled 'Out-of-Pocket' and Part 2 is entitled 'Coinsurance and Deductibles'. Each student plays both parts of the game. The 'Out-of-Pocket' scenario is initiated with the instructions:

> Players are self-employed contractors. Each player begins the simulation with an economic value of $10,000 and makes $3,000 quarterly. Each illness erodes the player's economic value through lost productivity ('Cost of Illness to Player') and each health care visit has the costs indicated by 'Cost of Care.' Players also have the option of undergoing an annual physical each year. Seeking care has an associated hourly time cost assessed at $10 per hour. Seeking care eliminates the cost of the illness to the player ('Cost of Illness to Player'). Each player makes $3000 a quarter and pays for health care out-of-pocket. Your goal is to maximise the 'Player's Cumulative Economic Value' at the close of the game.

The 'Coinsurance and Deductibles' scenario is initiated with the instructions:

> Players are self-employed contractors. Each player begins the simulation with an economic value of $10,000 and makes $3,000 quarterly. Each illness erodes the player's economic value through lost productivity ('Cost of Illness to Player') and each health care visit has the costs indicated by 'Cost of Care'. Seeking care has an associated hourly time cost assessed at $10 per hour. Players also have the option of undergoing an annual physical each year (to which the deductible and coinsurance apply). Seeking care eliminates the cost of the illness to the player ('Cost of Illness to Player'). Each player makes $2500 quarterly after a quarterly contribution of $500 to health insurance. The player pays for the care received until a $75 annual deductible is reached; thereafter, the health insurance pays 80% and the player pays 20%. Your goal is to maximise the 'Player's Cumulative Economic Value' at the close of the game.

Game play is designed to illustrate the differences in players' responses to seeking health care when paying out-of-pocket and when paying a premium to be covered by insurance with a coinsurance and deductible. The game is played using Blackboard®, the online learning system software whose use is pervasive throughout the United States system of higher education. The Test function in Blackboard® was used to improve an earlier version of the game which employed the static presentation of symptoms to players via spreadsheet (Kennedy *et al.*, 2005). Use of the Test function facilitated the

random presentation of symptoms at each of 12 health events to which players responded. This produced a more realistic scenario than the earlier version of the game which used a spreadsheet to present the symptoms experienced by the players all at once. Astute players could game the earlier version of the simulation by looking ahead to see which symptoms would occur in the future. If player responses illustrate moral hazard in the current version of the game, then this result is produced within a more realistic simulation environment. Game play via internet courseware also improved recording of player responses.

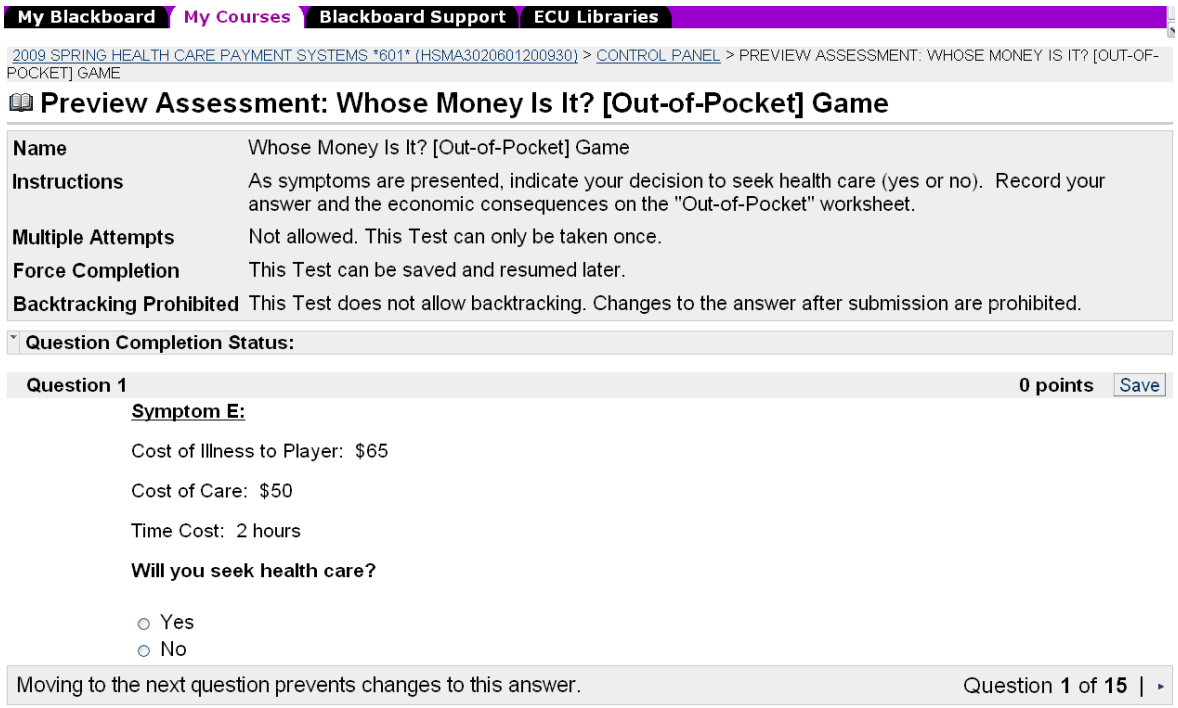
In the current version of the game, three years in the life of the player are simulated. Under both the 'Out-of-Pocket' and 'Coinsurance and Deductibles' scenarios, the player sequentially 'experiences' a quarterly health event presented randomly as one of 12 possible symptoms. Each presenting symptom is identified as *Symptom A* or *Symptom B, …* through to *Symptom L.*

Table 1: Presenting symptoms

Health Event	Cost of Illness to Player	Cost of Care	Time Cost in Hours
Symptom A	$75	$80	2
Symptom B	$65	$65	1
Symptom C	$110	$60	2
Symptom D	$95	$70	2
Symptom E	$65	$50	2
Symptom F	$200	$220	3
Symptom G	$55	$20	1
Symptom H	$25	$10	1
Symptom I	$35	$40	1
Symptom J	$70	$30	2
Symptom K	$25	$25	1
Symptom L	$40	$25	1

The player is also provided additional information characterising each symptom – the cost of illness to the player if health care is not sought, the cost of care if health care is sought, and the time cost in hours (valued at $10 an hour) if health care is sought. Figure 1 illustrates the typical presentation of a symptom by the courseware.

Figure 1: Presentation of a symptom by the courseware



At each presentation of a symptom, the player is asked 'Will you seek care?' and must submit an answer 'Yes' or 'No'. Depending upon the scenario involved, if the player decides to seek health care, the player pays for the cost of care 'out-of-pocket' or seeks care under the provisions of an insurance policy with a deductible and coinsurance. While submitting the decisions made to seek or forgo care as answers to a quiz presented via the courseware, the student also separately records his or her decision on a spreadsheet ledger to track both the economic consequences of that decision and the player's cumulative economic value as shown by Tables 2 and 3 for the respective scenarios.

Table 2: 'Out-of-pocket' worksheet

Year	Health Event	Cost of Illness to Player	Cost of Care	Time Cost (Hours)	Value of Time ($10/Hour)	Decision to seek health care? Answer "Yes" or "No"	Economic Value of Decision for Player	Quarterly Salary	Player's Cumulative Economic Value
									$10,000
	1							$3,000	
	2							$3,000	
1	Annual Physical								
	3							$3,000	
	4							$3,000	
	1							$3,000	
	2							$3,000	
2	Annual Physical								
	3							$3,000	
	4							$3,000	
	1							$3,000	
	2							$3,000	
3	Annual Physical								
	3							$3,000	
	4							$3,000	
								Total	

Table 3: 'Coinsurance and deductible' worksheet

Year	Health Event	Cost of Illness to Player	Cost of Care	Time Cost (Hours)	Value of Time ($10/Hour)	Coinsurance (20%)	Deductible Remaining	Seek health care? Answer "Yes" or "No"	Economic Value of Decision for Insurer	Economic Value of Decision for Player	Quarterly Salary Net of insurance	Player's Cumulative Economic Value
												$10,000
1	1						$75				$2,500	
	2										$2,500	
	Annual Physical											
	3										$2,500	
	4										$2,500	
2	1						$75				$2,500	
	2										$2,500	
	Annual Physical											
	3										$2,500	
	4										$2,500	
3	1						$75				$2,500	
	2										$2,500	
	Annual Physical											
	3										$2,500	
	4										$2,500	
											Total	

The goal for the player is to maximise cumulative economic value at the conclusion of the simulated three years of play. After submitting his or her decision via courseware and updating the spreadsheet ledger, the player responds to the next symptom presented by the software.

At the conclusion of the second, sixth, and tenth quarters, the midyear point of each year, players are presented with the opportunity to undergo an annual physical. The cost of the physical is $100 plus the hour time cost valued at $10. There is no short-term cost of illness for declining the physical, because in real life the decision to forgo a physical seldom results in short-term health consequences that erode economic value. As with the presentation of symptoms, the player is asked 'Will you seek care?' and must answer 'Yes' or 'No'. Including the annual physical facilitates the assessment of the value of physical examinations to students playing the game. Since selection of the physical always erodes the player's economic value while forgoing the physical never has an economic impact within the three years of game play, selection of the physical suggests that the player places future value on the physical examination.

Each scenario played concludes after three years of simulated play. Players' responses to the quarterly presentation of a health event (selected randomly with replacement as one of 12 symptoms) and to the three opportunities for an annual physical are recorded by the courseware. The player then submits the spreadsheet ledger, again via the Test function of the courseware, and responds to two questions: 'What strategy did you employ in deciding whether to seek care during the "Out-of-Pocket" game?' and 'What strategy did you employ in deciding whether to seek care during the "Coinsurance and Deductible" game?'

The game was played most recently in the online courseware environment as an assignment for two undergraduate health care payments systems classes (one face-to-face and one distance education) and for a graduate health care systems and problems class conducted as a distance course. The undergraduate classes were coeducational and respectively consisted of 52 face-to-face and 55 distance education junior and senior health services management and health information management programme majors. The graduate course was coeducational and consisted of 25 graduate students

pursuing a variety of health services management and health informatics concentrations and certificates at graduate level. Each class of students was graded on their participation in the game. Participation was worth 2.5% of the course grade for the undergraduate students and 3% of the course grade for the graduate students. As part of the institutional review process, each student in the three classes was provided with a copy of the draft article and requested to provide explicit permission to use his or her responses from game play. The responses of students not granting explicit permission were removed from analysis. Forty-one of 52 students (78.8%) from the face-to-face undergraduate health care payments systems class, 37 of 55 students (67.3%) from the distance education undergraduate health care payments systems class, and 20 of 25 students (80%) from the distance education graduate health care systems and problems class provided permission.

Students in each of the classes played both the 'Out-of-Pocket' and 'Coinsurance and Deductibles' scenarios after signing into the online Blackboard® environment at the time and place of their choosing. Game play was self-moderated, but each scenario could be played in as little as 10 to 15 minutes. Resources required to play the game were available online. The courseware could be accessed via computer and Internet browser. Student choices during both scenarios of game play were recorded by the quizzes supported by the courseware. Analysis of student responses was performed separately by downloading the response files in comma-delimited format and using Excel to perform the statistical computations.

3. Results

Game play clearly illustrates moral hazard. The one-tail paired-samples *t*-test for online game play by the face-to-face undergraduate health care payment systems students indicated that the students accessed care for symptoms significantly more frequently when paying premiums for insurance with a deductible and coinsurance (M = 7.76, SD = 2.48) than when paying for care out-of-pocket (M = 6.88, SD = 1.71) , $t(40) = 1.94$, $p = 0.030$. The one-tail paired-samples *t*-test for online game play by the distance education undergraduate health care payment systems students indicated that the students accessed care for symptoms significantly more frequently when paying premiums for insurance with a deductible and coinsurance (M = 8.16, SD = 2.43) than when paying for care out-of-pocket (M = 7.27, SD = 1.85) , $t(36) = 1.99$, $p = 0.027$. The one-tail paired-samples *t*-test for online game play by the distance education graduate health care systems and problems students indicated that the students accessed care for symptoms significantly more frequently when paying premiums for insurance with a deductible and coinsurance (M = 9.45, SD = 1.96) than when paying for care out-of-pocket (M = 7.05, SD = 1.28) , $t(19) = 4.71$, $p < 0.001$.

Similar results were indicated when the cost of care for symptoms treated were examined. The one-tail paired-samples *t*-test for online game play by the face-to-face undergraduate health care payment systems students indicated that the cost of care for symptoms treated was significantly more expensive for students when paying premiums for insurance with a deductible and coinsurance (M = \$365.00, SD = \$205.29) than for those paying for care out-of-pocket (M = \$290.37, SD = \$121.87), $t(40) = 1.91$, $p = .031$. The one-tail paired-samples *t*-test for online game play by the distance education undergraduate health care payment systems students indicated that the cost of care for symptoms treated was significantly more expensive for students when paying premiums for insurance with a deductible and coinsurance (M = \$411.49, SD = \$172.05) than for those paying for care out-of-pocket (M = \$302.16, SD = \$92.24), $t(36) = 3.32$, $p = 0.001$. The one-tail paired-samples *t*-test for online game play by the distance education graduate health care systems and problems students indicated that the cost of care for symptoms treated was significantly more expensive for students when paying premiums for insurance with a deductible and coinsurance (M = \$486.25, SD = \$208.04) than for those paying for care out-of-pocket (M = \$288.00, SD = \$104.65), $t(19) = 3.93$, $p < 0.001$.

Between the second and third quarter of each year's game play, students also made a choice about whether to undergo an annual physical. Selecting the physical cost $100 for cost of care and $10 in time costs; there were no short-term costs of illness for forgoing the physical. Some students chose to undergo an annual physical even though it eroded their economic value. Online game play by the face-to-face undergraduate health care payment systems students indicated that players chose to have a physical on average 1.61 times in a three-year period when paying premiums on insurance with a deductible and coinsurance compared to 1.34 times when paying out-of-pocket. A one-tail paired-samples *t*-test indicated that the difference was significant (M = 1.61, SD = 1.12 compared to M = 1.34, SD = 1.09), $t(40) = 1.76$, $p = 0.043$. Since the cost of the physical is constant at $100, compared to the earlier presentation of symptoms, the dollar relationship to the times physicals are selected is constant as well. Distance education students in the undergraduate health care payment systems class opted for a physical more frequently and a one-tail paired-samples *t*-test indicated that the difference in selection when players pay premiums for insurance with a deductible and coinsurance (M = 2.30, SD = 1.05) than when paying out-of-pocket (M = 1.62, SD = 1.23), was significant, $t(36) = 3.29$, $p = 0.001$. Distance education students in the graduate health care systems and problems class also illustrated moral hazard when choosing to undergo a physical. A one-tail paired-samples *t*-test indicated that the difference in selection when players pay premiums for insurance with a deductible and coinsurance (M = 2.40, SD = 0.94) than for those paying for care out-of-pocket (M = 1.05, SD = 1.28), was significant, $t(19) = 4.76$, $p < 0.001$.

Results of game play are summarised in Table 4.

Table 4: Results

Measure	Class	Payment Status	M	SD	N	*t*	*p (one-tailed)*
Accessed Care	Face-to-Face Undergraduate Health Payment Systems	Insurance	7.76	2.48	41	1.94	0.030
		Out-of Pocket	6.88	1.71			
	Distance Education Undergraduate Health Payment Systems	Insurance	8.16	2.43	37	1.99	0.027
		Out-of Pocket	7.27	1.85			
	Distance Education Graduate Health Care Systems and Problems	Insurance	9.45	1.96	20	4.71	< 0.001
		Out-of Pocket	7.05	1.28			
Cost of Care	Face-to-Face Undergraduate Health Payment Systems	Insurance	$365.00	$205.29	41	1.91	0.031
		Out-of Pocket	$290.37	$121.87			
	Distance Education Undergraduate Health Payment Systems	Insurance	$411.49	$172.05	37	3.32	0.001
		Out-of Pocket	$302.16	$92.24			
	Distance Education Graduate Health Care Systems and Problems	Insurance	$486.25	$208.04	20	3.93	< 0.001
		Out-of Pocket	$288.00	$104.65			
Chose Physical	Face-to-Face Undergraduate	Insurance	1.61	1.12	41	1.76	0.043
		Out-of Pocket	1.34	1.09			
	Distance Education Undergraduate	Insurance	2.30	1.05	37	3.29	0.001
		Out-of Pocket	1.62	1.23			
	Distance Education Graduate Health Care Systems and Problems	Insurance	2.40	0.94	20	4.76	< 0.001
		Out-of Pocket	1.05	1.28			

4. Discussion

When presented with symptoms, moral hazard was illustrated by the students of all three classes. They sought care more frequently when insured and the cost of care sought was greater. Student completion of the spreadsheet ledgers forced students to consider the economic consequences of their choices and provided the instructor with the opportunity to check whether students understood how to

calculate premium costs, out-of-pocket costs, and deductible and coinsurance cost-shares. Responses to the questions 'What strategy did you employ in deciding whether to seek care during the "Out-of-Pocket" game?' and 'What strategy did you employ in deciding whether to seek care during the "Coinsurance and Deductibles" game?' served as the catalyst for discussion in the face-to-face undergraduate health care payment systems class which was recorded and viewed by the distance education class. The students in the graduate health care systems and problems class discussed their game play in an online chat session. Interpretation by the author of the students' expressed strategies with and without insurance was somewhat subjective, but led to the following judgment of student responses which is summarized by Table 5. Some students were clearly able to articulate the optimal strategy for accumulating economic value when playing the 'Out-of-Pocket' scenario – Choose Care when Cost of Care + Time Cost < Cost of Illness. Responses ranged from 35% for the face-to-face undergraduate health care payment systems students to approximately 40% for both the distance education undergraduate health care payment systems students and the graduate health care systems and problems students taught at a distance. A smaller percentage (2.50% for the face-to-face undergraduate health care payment systems students, 8.11% for the distance education undergraduate health care payment systems students, and 13.33% of the distance education graduate health care systems students) professed confusion or having figured a strategy out only after completing game play. The remainder of the students for each class was only able to articulate a partial strategy.

The best strategy was less clear during the 'Coinsurance and Deductible' scenario. For those players able to articulate a distinct, rational strategy, two strategies emerged: 1) Choose Care when Cost of Care + Time Cost < Cost of Illness while monitoring the effect on cost-sharing once the deductible was consumed and 2) Consume the deductible as quickly as possible to gain the cost-sharing advantage of insurance. 15.00% of the face-to-face undergraduate health care payment systems students, 21.62% of the distance education undergraduate health care payment systems students, and 6.67% of the graduate health care systems and problems students taught at a distance clearly articulated the strategy Choose Care when Cost of Care + Time Cost < Cost of Illness while monitoring the effect on cost-sharing once the deductible was consumed. 20.00% of the face-to-face undergraduate health care payment systems students, 21.62% of the distance education undergraduate health care payment systems students, and 26.67% of the graduate health care systems and problems students taught at a distance chose to seek care and/or undergo a physical in order to consume the deductible for insurance to come into play. A smaller percentage (5.00% for the face-to-face undergraduate health care payment systems students, 2.70% for the distance education undergraduate health care payment systems students, and 6.67% of the graduate health care systems and problems students taught at a distance) professed confusion or having figured a strategy out only after completing game play. The remainder of the students for each class were only able to articulate a partial strategy.

One phenomenon of game play was highlighted by further discussion. Once the deductible was consumed, insurance coverage allowed care to be sought at 80% discount. Even so, in the absence of catastrophic loss not represented by the symptoms presented during game play, this insurance advantage was not sufficient to overcome the cost of the premiums. Noting that fact and asking a question about the value of insurance in those circumstances facilitated further discussion about risk aversion and the role of insurance in mitigating the risk of catastrophic loss.

Table 5: Strategy

Class	Strategy	Out-of-Pocket	Copy and Deductible
Face-to-face undergraduate health payment systems[1]	Choose care when cost of care + time cost < cost of illness	35.00% (14)	15.00% (6)
	Partial strategy articulated	62.50% (25)	60.00% (24)
	Consume deductible	–	20.00% (8)
	Incorrect strategy/no strategy	2.50% (1)	5.00% (2)
Distance education undergraduate health payment systems	Choose care when cost of care + time cost < cost of illness	40.54% (15)	21.62% (8)
	Partial strategy articulated	51.35% (19)	54.05% (20)
	Consume deductible	–	21.62% (8)
	Incorrect strategy/no strategy	8.11% (3)	2.70% (1)
Distance education graduate health care systems and problems[2]	Choose care when cost of care + time cost < cost of illness	40.00% (6)	6.67% (1)
	Partial strategy articulated	46.67% (7)	60.00% (9)
	Consume deductible	–	26.67% (4)
	Incorrect strategy/no strategy	13.33% (2)	6.67% (1)

The opportunity to choose to undergo an annual physical was an interesting nuance of game play. Students from all three classes chose to undergo at least one physical on average when paying out-of-pocket within the three years represented by game play. This choice was made even though the physical clearly reduced short-term cumulative economic value. Responses by some of the students indicated that they were willing to accept this loss because of the benefits perceived from having a physical. Although not every student choosing to receive a physical when paying out-of-pocket discussed their choice as a part of their strategy, of the 24 students from the three classes in total who explicitly stated that they had opted to receive at least one physical during the three-year period of game play, 18 commented that physicals had preventive value, were important, or were perceived to influence the outcome of the game. This suggests that some students regarded the physical as a kind of insurance. Although the probability of discovering imminent, catastrophic illness as the result of a physical is low, risk aversion provides an incentive to manage risk through the periodic screening offered by a physical even if payment is made out-of-pocket. If risk aversion contributes to the purchase of a physical out-of-pocket, moral hazard suggests that the availability of insurance should contribute to the decision to purchase additional physicals as was demonstrated by game play. Indeed, the students in all three classes chose to undergo a physical significantly more frequently when paying for insurance with a deductible and coinsurance. Although anticipated, this is not a trivial result. Another possible response illustrative of moral hazard was suggested by Helwege (1996, p. 60) while describing a policy exercise: 'The patient may engage in a suboptimal level of preventive care, knowing that an insurer will bear a significant portion of an acute health crisis.' In this exercise, some students responded that they made the choice to purchase additional physicals because of the benefits perceived from having a physical and others as a means of consuming the deductible more quickly to bring the benefits of insurance into play.

[1] One response not available for analysis

[2] Five responses not available for analysis

5. Conclusions

Playing the two scenarios of this moral hazard game engages students in deciding whether to seek care when presented with symptoms after considering the economic impact of their decisions. The concept of moral hazard is illustrated by game play rather than simply cited during lecture, and students are actually able to see during the after-action briefing how they or their peers tend to seek care more frequently and at greater cost when insured.

The use of online courseware in both face-to-face and distance classes produced additional advantages:

1. The courseware was familiar to both students and faculty.
2. Use of the online courseware facilitated a dynamic environment. Symptoms and associated costs were presented randomly at scheduled intervals (in this case, quarterly). When a previous version of the game was played as a static spreadsheet game, all of the symptoms and costs were presented at once. Even though players were supposed to make decisions to seek care iteratively, the ability to 'look ahead' to view upcoming symptoms allowed students to game their selections in the earlier version.
3. Functions inherent to the courseware required no specific programming skills by the instructor.
4. Administration of the game could be packaged effectively within the courseware. Preview materials to be read prior to game play, the game itself, and debriefing/after-action materials all could be packaged within the same folder.
5. Responses were conveniently recorded for retrieval and analysis. Although arranging the data so that Excel could perform the statistical computations after the response files were downloaded from the courseware was somewhat tedious, the procedure was still an improvement over the previous version of the game.

Finally, play of the game facilitated further classroom discussion about risk aversion and insurance theory.

Author's note

The author thanks the editor and two anonymous referees for constructive feedback. This game is available to other instructors teaching with Blackboard®. Contact the author by email [kennedym@ecu.edu] or telephone [+1 (252) 744-6182] for assistance in importing the necessary files.

References

Brown, D. G. (2001). 'Hybrid courses are best', *Syllabus,* vol. 15(1), p. 22.

Feldstein, P.J. (2005). *Health Care Economics* (6th edn), Florence, KY: Delmar Cengage Learning.

Folland, S., Goodman, A.C. and Stano, M. (2010). *The Economics of Health and Health Care* (6th edn), Upper Saddle River, New Jersey: Prentice Hall.

Foshay, W.R. (2002). 'Distance learning: Is that all there is, my friend?' *The Quarterly Review of Distance Education,* vol. 3(1), pp. 65–74.

Friedman, D. and Sunder, S. (1994). *Experimental Methods: A Primer for Economists,* New York, NY: Cambridge University Press.

Getzen, T.E. (2010). *Health Economics and Financing* (4th edn), Hoboken, NJ: John Wiley & Sons, Inc.

Helwege, A. (1996). 'Preventive versus curative medicine: A policy exercise for the classroom', *Journal of Economic Education,* vol. 27(1), pp. 59–71.

Henderson, J. W. (2009). *Health Economics and Policy* (4th edn), Mason, Ohio: South-Western Cengage Learning.

Jonassen, D.H. (2002). 'Engaging and supporting problem solving in online learning', *The Quarterly Review of Distance Education,* vol. 3(1), pp. 1–13.

Kagel, J.H. and Roth, A.E. (eds) (1995). *The Handbook of Experimental Economics,* Princeton, NJ: Princeton University Press.

Kennedy, M.H., Unruh, L. and Kerr, B.J. (2005). 'Whose money is it?: An economic game on the concept of moral hazard'. Available at SSRN: http://ssrn.com/abstract=846366

Larreamendy-Joerns, J. and Leinhardt, G. (2006). 'Going the distance with online education', *Review of Educational Research,* vol. 76(4), pp. 567–605.

Morrisey, M.A. (2007). *Health Insurance*, Chicago, IL: Health Administration Press.

Ortmann, A. and Colander, C. (1997). 'Teaching tools: A simple principal-agent experiment for the classroom', *Economic Inquiry,* vol. 35(2), pp. 443–50.

Vrasidas, C. (2004). 'Issues of pedagogy and design in e-learning systems', in *ACM Symposium on Online Learning,* pp. 911–15). New York, NY: ACM Press.

Author Biography

Dr. Michael H. Kennedy is an Associate Professor and Director of the Health Services Management Program in the College of Allied Health Sciences at East Carolina University, Greenville, NC, USA. He has 35 years experience in teaching and health services administration that have been divided between academic positions and operational assignments in the United States' military health system. Dr. Kennedy earned his doctorate in Decision Sciences and Engineering Systems from Rensselaer Polytechnic Institute.

Contact details

Michael H. Kennedy
Associate Professor and Director
Health Services Management Program
Department of Health Services and Information Management
East Carolina University, USA
Tel: +1 (252) 744 6182
Email: kennedym@ecu.edu