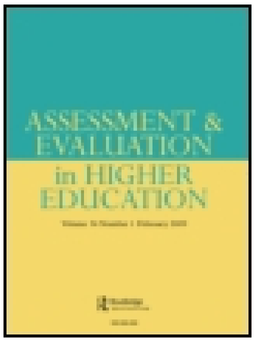

Assessment & Evaluation in Higher Education

ISSN: 0260-2938 (Print) 1469-297X (Online) Journal homepage: https://www.tandfonline.com/loi/caeh20

Allowing course compensation in higher education: a latent class regression analysis to evaluate performance on a follow-up course

Iris E. Yocarini, Samantha Bouwmeester, Guus Smeets & Lidia R. Arends

To cite this article: Iris E. Yocarini, Samantha Bouwmeester, Guus Smeets & Lidia R. Arends (2019): Allowing course compensation in higher education: a latent class regression analysis to evaluate performance on a follow-up course, Assessment & Evaluation in Higher Education, DOI: 10.1080/02602938.2019.1693494

To link to this article: https://doi.org/10.1080/02602938.2019.1693494

Published online: 29 Nov 2019.

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
https://doi.org/10.1080/02602938.2019.1693494

Routledge
Taylor & Francis Group

OPEN ACCESS

Check for updates

Allowing course compensation in higher education: a latent class regression analysis to evaluate performance on a follow-up course

Iris E. Yocarini[a,b], Samantha Bouwmeester[a], Guus Smeets[a] and Lidia R. Arends[a]

[a]Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands; [b]Leiden Institute of Advanced Computer Science, Leiden University, Leiden, The Netherlands

ABSTRACT

In this study, the consequences of allowing course compensation in a higher education academic dismissal policy are evaluated by examining performance on a second-year follow-up (i.e. sequel) course that builds on material from a first-year precursor course. Up to now, differences in the consequences of compensation on student performance across groups of students who portray different unobserved study processes were not considered. In this study we used a latent class regression model to distinguish latent groups of students. Data from two undergraduate curricula were used and latent classes were formed based on similar patterns in averages, variability in grades, the number of compensated courses, and the number of retakes in the first year. Results show that students can be distinguished by three latent classes. Although the first-year precursor course is compensated in each of these latent classes, low performance on the precursor course results in low performance on the second-year sequel course for psychology students who belong to a class in which the average across first-year courses is low and the average number of compensated courses and retakes are high. For these students, compensation on a precursor course seems more likely to relate to insufficient performance on a sequel course.

KEYWORDS

Compensation; higher education; latent class regression; academic performance

Student success in higher education is an important issue. This is underlined by the goal set in the Europe 2020 strategy to have at least 40% of 30–34-year-olds complete higher education. Reducing student dropout and increasing study completion rates is one of the main strategies to improve student success (Vossensteyn et al. 2015). One successful intervention, which we focus on in this study, is the use of an academic dismissal policy, that is a performance-based selection mechanism through which students may be dismissed from an academic program (in The Netherlands referred to as the binding study advice; Sneyers and De Witte 2018). In an academic dismissal policy, students' progress is evaluated for example after the first year of the bachelor to assess whether the requirements to continue their studies are met. In making this decision, different decision rules may be applied by higher education institutions. Traditionally, a conjunctive decision rule is applied, in which students either pass or fail an individual course. Here, study

CONTACT Iris Yocarini i.e.yocarini@liacs.leidenuniv.nl

credits are assigned to individual grades. Alternatively, a compensatory decision rule may be applied in which students are allowed to compensate, within boundaries, a low score on one course with a high score on another course. In this situation, students receive study credits based on their average score. In the present study, the latter approach is examined as we aim to evaluate performance of students who are allowed to compensate courses in the academic dismissal policy.

Allowing course compensation

Different reasons may motivate the implementation of a compensatory decision rule instead of a conjunctive decision rule. First, one might allow compensation to improve students' grade goals and their motivation to perform well on tests. In a compensatory system it pays to get a grade that is as high as possible as the average grade serves as the selection instrument, instead of just passing a test. Furthermore, one might implement compensation with the intention to decrease students' procrastination by limiting the number of retakes allowed and the challenges associated with retakes (Pell, Boursicot, and Roberts 2009; Arnold 2017). When compensation is allowed, a failing grade on a course does not necessarily need to be retaken and subsequent study delay might be decreased. Third, students are trained for a profession that is compensatory by nature. In a job, an employee may compensate his or her lacking (relatively, above a certain required minimum level) in one area by outshining others in another area of expertise (Rekveld and Starren 1994). Fourth, the argument that the average grade is more reliable than individual test scores and consequently guards against making incorrect decisions in the academic dismissal policy, may motivate the implementation of a compensatory rule (De Gruijter 2008).

Whilst there may be several reasons to implement compensation, there is also critique. Opponents argue that compensation might result in unfavourable study behaviour that might result in lower academic performance. When compensation is allowed, students may make certain strategic study choices in terms of their resource allocation (such as in time and effort; see e.g. van Naerssen 1970) that might hamper their academic performance. Certain resource allocation strategies, such as for example, focusing more on easier courses and less on difficult ones, might possibly create problematic hiatuses (i.e. gaps) in students' knowledge, thereby decreasing educational quality (Arnold 2011). Specifically, this concern applies to the situation in which courses accumulate on knowledge obtained in previous courses (so-called sequel courses). In this way, allowed to compensate, students might not obtain sufficient knowledge to perform well on a sequel course and possibly graduate with hiatuses.

To prevent such undesired situations, educational programs are mostly designed accordingly, and for example define clusters of courses in which compensation is allowed (Rekveld and Starren 1994). By forming clusters based on for example the course content or the course difficulty level, or by giving difficult courses more weight, undesired strategic study behaviour potentially causing graduates to have hiatuses in their knowledge can be avoided. Consequently, decision rules in an educational context are rarely fully compensatory (Douglas and Mislevy 2010). Rather, there are some conjunctive aspects included in which a minimum level of performance is required. Previous research, in which the argument that the average grade is more reliable was evaluated, shows that the required minimum grade is also important for the accuracy of the compensatory rule (Yocarini et al. 2018). Furthermore, the results of this study show that the accuracy of a compensatory decision rule relative to a conjunctive decision rule depends on the test reliabilities, correlation between tests, and the number of resits allowed. As such, choosing a specific decision rule should involve the evaluation of the decision accuracy as well as the characteristics of the tests. Hereby, false negatives, those students who failed but are truly competent enough to pass, are more prominent in conjunctive decision rules than false positives, students who passed but are not sufficiently skilled yet. For compensatory decision rules the reverse is true and more false positives occur compared to false negatives. The discussion of whether to allow course compensation is therefore also a discussion of weighing the disadvantages of false negatives over false positives and vice versa.

With the increased focus on student success in higher education, the debate on allowing compensation in an academic dismissal policy has gained more attention as well (see e.g. Smits, Kelderman, and Hoeksma 2015). Several studies have focused on the evaluation of compensatory decision rules. For example, Arnold (2011) evaluated the consequences in an economics bachelor program and showed that whether the compensated course grade was obtained after one or multiple tries (i.e. was retaken) was important for later performance; that is, the number of retakes were negatively related to performance on the sequel course. Up to now the debate has not yet touched upon the question of whether the discussion of the consequences of compensation (e.g. possible hiatuses in knowledge) applies to each specific group of students. Although this point was raised by Smits, Kelderman, and Hoeksma (2015), no study has evaluated this. The higher education student population is diverse, containing students with varying levels of cognitive abilities who may portray different study strategies. Students' grades and choices to compensate or retake courses in a curriculum may therefore vary.

The present study

The purpose of this study is to take into account unobserved study processes for students who are allowed to compensate courses. In this way, our study extends previous studies that have studied the relation between performance on a first-year precursor and second-year sequel course. To our knowledge, there is no published study that evaluated differences between groups of students that show similar study processes. As these groups have yet to be detected, and consequently are not yet observed, the existence of these groups could be explored by means of a latent class model. In a latent class model, it can be evaluated whether groups of students exist who share the same pattern of values on the observed variables. These groups are called latent because they are not a priori defined by a manifest variable.

Variables that were used to distinguish these latent classes are the first-year average grade, the variation in first-year grades, the number of courses that were compensated, and the number of courses that were retaken. Courses are qualified as compensated when the course grade (for example, 5.0, on a 1-10-point scale, as is common in Dutch higher education) is below the required average grade set in the compensatory decision rule (e.g. 6.0). These variables were selected as these are expected to be able to make a distinction between groups of students who make different (unobserved) choices with regards to their study resource allocation in the first year of the bachelor, and which groups may display different relations between first-year precursor and second-year sequel course performance. Although students with very high and very low cognitive abilities will probably perform similarly under different decision rules (i.e. pass or fail in either situation), students with average abilities just above the cut-score are of most interest in the compensation discussion.

Here, two students (let's call them Ann and Peter) with similar average grades might have obtained this average through a different pattern of grades. Where Ann might have obtained her average grade by consistently scoring around this average, Peter might have high variability in his grades. Study choices as to which course to compensate or retake, and to alter study resource allocation accordingly, might be more useful to Peter compared to Ann. With more variation in his grades, Peter might choose to put extra effort in the courses for which he expects to perform well and to neglect courses for which he believes he'll get a low grade anyway. Consequently, knowledge gaps are more likely for students like Peter. Therefore, performance on a second-year sequel course that builds on a 'neglected' precursor course is more likely to be low for Peter compared to Ann. By compensating this course, it might be that Peter lacks knowledge of the first-year material required in the second-year sequel course, resulting in lower academic performance in the second-year compared to Ann. Up to now, the consequences of compensation have been studied as being similar for Ann and Peter by not making a distinction in latent student groups.

The overall aim of this study is to evaluate the consequences of course compensation by evaluating performance on a second-year sequel course across different latent student groups. Specifically, the relation between the first-year precursor course and second-year sequel course performance is allowed to vary across latent student groups who are characterized by a similar pattern in first-year grades, variability in their first-year grades, the number of compensated first-year courses, and the number of retakes in the first year. For this purpose, a latent class regression model (Wedel & DeSarbo, 1994) is applied on data from a Dutch psychology university bachelor in which compensation is allowed within boundaries in the first year of the bachelor. As a second purpose, the generalizability of the results is assessed by replicating the analyses on data from a Dutch law university bachelor program in which course compensation is also allowed. Here, a latent class model has the advantage that groups of students can be formed that show similar grade characteristics and study choices without prior assumptions about the specific formation of these groups (e.g. in terms of the number of classes or class sizes).

Method

Sample

Test scores from students' first- and second-year courses in a psychology bachelor program at a Dutch university were used. These data were obtained from a large database of a Dutch university. Specifically, cohorts were included in which compensation was allowed and in which a sequel second-year course was present. For this selection, the content of the courses was considered by consulting the examination regulations and course descriptions, and course coordinators or program executives. Overall, only students who passed the academic dismissal policy requirements were included. This implies that each included student obtained a grade on each first-year course (i.e. no missing values were allowed). Dutch students are graded on a scale of 1 to 10 with 5.5 serving as the cut-score for a passing grade. Compared to American grading scales, a grade of 8.Five or higher corresponds to an 'A+', a grade of eight to an 'A', a grade of seven to a 'B+', a grade of 6.5 to a 'B', a grade of a six to a 'C', a grade of 5.5 to a 'D', and a grade of five or lower to a 'F' (Nuffic, 2009). Following these selection criteria, the cohorts 2011–2015 were selected, including 1077 psychology students. These students were required to score 6.0 on average (rounded from 5.95 on a 1-10-point scale) over eight courses with a minimum required grade of 4.0 on each individual course. Of these eight tests only two were allowed to be retaken once. Overall, one course combination existed in which the second-year course very clearly and explicitly built on first-year material for these psychology cohorts and the first-year course was compensated relatively often: Statistics I in the first year and Statistics II in the second year.

To assess the generalizability of our findings, students' grades from the law bachelor were selected to replicate the analyses. If one would expect similarities in latent classes across study programs, these similarities are expected to be most pronounced in a study program such as the law bachelor that is most similar to the psychology program with respect to the organisation (i.e. eight consecutive courses that each have a similar number of course credits), didactic approach (i.e. problem-based learning), size, and academic dismissal policy decision rule. Following a similar exclusion procedure, the cohorts 2012–2015 were selected, including 1120 law students who were required to score 6.0 on average (unrounded) over eight courses with a minimum required grade of 4.5 on each individual course. Similarly, two out of these eight tests were allowed to be retaken once. The first-year precursor and second-year sequel course combination in which the first-year course was compensated most often was selected: Introduction

Table 1. Descriptive statistics continuous variables.

Study program	Course	Year	Variable	Mean	Median	SD	Min	Max
Psychology		1	Yearly average[a]	6.77	6.60	0.65	5.95	9.25
		1	Yearly *SD*	0.89	0.88	0.24	0.28	1.67
	Statistics I	1	Course grade	6.40	6.40	1.28	4	10
	Statistics II	2	Course grade[b]	6.66	6.80	1.49	1	10
Law		1	Yearly average[a]	6.82	6.75	0.61	6	9
		1	Yearly *SD*	0.80	0.79	0.22	0	1.51
	Intro constitutional and administrative law	1	Course grade	6.54	7	1.03	5	10
	Constitutional law	2	Course grade[c]	6.26	6	1.08	3	10

[a]The first year average was only computed for students who received a grade for all first-year courses in their study program.
[b]Note that second-year courses do not have the requirements in the academic dismissal policy as in the first year, so grades run from 1 to 10, *NA* = 77.
[c]*NA* = 192.

to constitutional and administrative law in the first year and Constitutional law in the second year (see Tables 1 and 2).

As shown in Table 1 the yearly averages of both curricula are distributed quite similarly, with the distinction that grades in law are rounded whereas grades in psychology are not. Furthermore, compared to the psychology data, the precursor and sequel course grade in law have less variation and a smaller range as shown by the min and max grade. Table 2 shows that in the psychology curriculum students compensated more courses, whereas students in the law curriculum seem to have retaken more courses. Focusing on the first-year precursor course specifically shows that psychology students compensated this precursor course more often. These differences across the study programs might be due to differences in the course combinations, such that the sequel course in the psychology curriculum more extensively builds on the precursor course than in the course combination in the law program. Also, grades on the psychology courses might be lower because the courses might be more difficult within the curriculum than the courses evaluated in the law program.

Statistical analyses

To assess the relation between the grade on the first-year precursor course and the second-year sequel course grade across different groups of students a latent class regression model is applied. In a latent class regression model the dependent variable is class dependent. In our study this model implies that the grade on a second-year sequel course may vary depending on the specific latent class a student belongs to. Besides the dependent variable, grade on a second-year sequel course, two kinds of independent variables are included. First, there are predictor variables that may explain variance in the dependent variable, as is common in traditional regression analyses. Here, the grade on the first-year precursor course is included as a predictor variable to assess the relation between the first-year course grade with the second-year course grade. Second, independent variables (referred to as covariates) that may explain the existence of different latent classes are included. In the current study, four covariates are included, namely the yearly average, variation in first-year grades, the number of compensated courses in the first year (i.e. the number of course grades below the required minimum grade), and the number of resits in the first year. Based on students' values on these four covariate variables, the latent classes are formed such that students belonging to the same latent class are characterized by similar covariate values. Because the second-year sequel course grades depend on these classes as well, the latent classes in this latent class regression analysis are formed such that they depend both on the first-year grade patterns as well as the grade on a second-year sequel course.

Table 2. Descriptive statistics categorical variables.

Study program	Course	Year	Variable	0	1	2	3	4	5
Psychology		1	Number of compensated courses[a] (< 6.0)	25.5% (275)	22.9% (247)	20.3% (219)	20.3%(219)	8.6% (93)	2.2% (24)
		1	Number of resits	71.7% (772)	11% (119)	17.3%(186)			
	Statistics I	1	Course compensated	62.1% (669)	37.9% (408)				
		1	Course retaken	90.3% (972)	9.7% (105)				
Law		1	Number of compensated courses[a]	52.3% (586)	27.1% (304)	16.8% (188)	3.4% (38)	0.4% (4)	
		1	Number of resits	50.7% (568)	34.6% (387)	14.7% (165)			
	Intro constitutional and administrative law	1	Course compensated	83.2% (932)	16.8% (188)				
		1	Course retaken	89% (995)	11.2% (125)				

[a]Compensated courses are courses on which the grade was below the required average of 6.0. This thus differs from the number of insufficient grades (below the Dutch cut-off of 5.5).

Table 3. Validation information criteria and classification errors for different LC models for psychology.

Model	*LL*[a]	BIC	AIC	AIC(3)	Number of parameters	Proportion of classification errors
1- class	−1732.44	3485.61	3470.87	3473.87	3	0
2- class	−1632.06	3340.16	3286.12	3297.12	11	0.12
3- class	−1605.28	3341.91	3248.57	3267.57	19	0.2
4- class	−1616.03	3418.71	3286.07	3313.07	27	0.25
5- class	−1611.22	3464.39	3292.45	3327.45	35	0.22
6- class	−1615.40	3528.05	3316.80	3359.80	43	0.22

[a]*LL* = Log-Likelihood. Note that the Log-Likelihood slightly increases at the 4- and 6-class model, this is possible because of the holdout validation procedure used to estimate these values.

In a latent class regression analysis, the regression weights for both kinds of independent variables (predictors and covariates) are estimated simultaneously by an optimisation procedure for the complete set of parameters in the model. To select the number of classes, different models with increasing numbers of classes were fitted to the data. Consequently, the fit of these models was compared to select the best fitting model. The latent class regression was performed using Latent GOLD 5.0 (LG; Vermunt & Magidson, 2013: the syntax may be obtained upon contacting the corresponding author). After fitting the latent class regression model to the psychology data, the latent class model was further validated by performing the same analysis on the law data. In this way, the generalizability of our findings across study programs was assessed.

Results

Latent class regression analysis

Several latent class regression models were fitted to the psychology data, with an increasing number of classes. The validation fit statistics and proportion of classification errors for the latent class models are displayed in Table 3.

The number of classes were determined using various information criteria: the Akaike information criterion (AIC), the Akaike information criterion 3 (AIC3), and the Bayesian information criterion (BIC). Each of these indices applies different penalties on the log-likelihood statistic for the number of model parameters, sample size, or both, and may therefore point towards different best fitting models. First, the BIC values in Table 3 are lowest for the two-class model and the AIC and AIC(3) values are lowest for the three-class model. As shown, the proportion of classification errors, which indicate how distinct the latent classes are, are higher for the three-class model. These classifications are high because some students are not easily classified in one of the three classes within the latent class model. Upon assessing the classification errors per class, this seems mostly true for students classified in class one and two. The class sizes show that in both models there is one relatively larger class and one or two smaller classes. Where the two-class model consists of one class that has a high first-year average and one that has a low first-year average, the three-class model makes an additional distinction resulting in two classes with low and average first-year averages. Because this additional distinction adds valuable information for our latent class regression analysis, the three-class model was selected.

The parameter estimates of the covariates in the latent class regression, as shown in Table 4, can be tested to see whether the influence of the covariate has a significant influence on the classes. The results for these tests showed that the yearly average, Wald statistic (2) = 29.16, $p <$ .001, the yearly number of compensated courses, Wald statistic (2) = 8.30, p = .016, and the yearly number of retaken tests, Wald statistic (2) = 13.69, $p <$ .001, were significantly different across the latent classes. As shown in Table 4, the first class can be interpreted as the students with a low performance on average (6.2), a high average number of compensated courses (about three) and a high average number of retaken tests (more than one). About a quarter of the

Table 4. Descriptive statistics for the three-class model for psychology.

Variable		Class 1	Class 2	Class 3
First-year average[a]	Average	6.20	6.69	7.61
	SD	0.20	0.43	0.62
	Max	7	8	9
First-year standard deviation[b]	Average	0.88	0.91	0.86
	SD	0.22	0.23	0.25
Number of compensations[c]	Average	2.86	1.61	0.63
	SD	1.07	1.22	1.09
Number of retaken tests[d]	Average	1.35	0.21	0.09
	SD	0.87	0.51	0.33
First-year precursor course grade[e]	Average	5.64	6.35	7.32
	SD	1.03	1.14	1.25
	Max	9	9	10
Second-year sequel course grade	Average	4.98	6.68	8.47
	SD	1.26	0.90	0.61
	Min	1	4	7
	Max	9	9	10
Class size		0.23	0.56	0.21
N[f]		234	563	208

[a]The minimum first-year average is six for all classes.
[b]The range of the standard deviation in first-year grades is similar across classes: from 0 to 2.
[c]The number of compensated courses ranges between 0 and five for each class.
[d]The number of retaken tests ranges between 0 and two for each class.
[e]The minimum grade is four across all classes.
[f]Sample size here is smaller as reported in the Results section as there were 72 missing values on the second-year grade.

students are classified in class one. The second class, about half of the sample, are students with average performance levels (6.7), a moderate average number of compensated courses (about one or two) and a low number of retaken tests on average (mostly none or one). Finally, the third class, about a fifth of the sample, consists of the high performing students (average grade of 7.6), who have a low number of compensated courses on average (mostly none or one) and no retakes on average.

Subsequently, the class-dependent relation between the first-year precursor grade on the second-year sequel course grade was evaluated. Here, a Wald test indicated the relation of the predictor with the second-year sequel course grade to be significant: Wald statistic (3) $=$ 135.09, $p < .001$. When the precursor course grade was high, the grade on the second-year sequel course was high as well. Yet, a Wald test comparing the parameters across classes was not significant: Wald statistic (2) $=$ 0.20, $p = .90$, indicating that the parameters did not significantly differ across classes. This implies that the positive relation that is found between the first-year precursor and second-year sequel grade did not vary statistically significant across different latent classes. Furthermore, the variances of the dependent variable, the second-year sequel course grades, were significantly different across the three classes: Wald statistic (2) $=$ 42.05, $p <$ .001, showing that the variability in second-year sequel course grades differs across the latent classes. As shown in Table 4, variation was highest in the low performing class (1.26 for class one), lower for the moderate performing class (0.90 for class two) and smallest for the high performing class (0.61 for class three). These different variations per class influence the significance of the difference in the parameters of the predictors across the classes. Therefore, it is also important to evaluate the average performance on the first-year precursor and second-year sequel course across classes in addition to the parameter values of the latent class regression analysis.

The lower half of Table 4 shows these estimates. Importantly, for students in the first class, the average grade on the precursor grade was 5.64 and therefore just sufficient at the Dutch cut-off score of 5.5 yet below the required average grade of 6.0 in the first-year compensatory decision rule. For these students, the average grade on the second-year sequel course was lower and insufficient at 4.98. For students in the second class, the average grade on the first-year course was 6.35 and therefore sufficient and above the required average of 6.0. For this group, the average performance on the second-year sequel course was 6.68. Finally, for students in the third class the average grade on the first-year course was 7.32 and the average grade on the second-year course was high as well at 8.47. The ranges of the first-year and second-year grades show, as defined by the minimum and maximum values in Table 4, that while in every class (some) students compensated the precursor course grade, the second-year sequel course was only compensated by (some) students from the first and second class. Taken together, these results seem to suggest that students whose first-year performance is low and consequently compensate and/or fail the precursor course in the first-year have a higher likelihood to perform low on the sequel course.

Latent class regression analysis law curriculum

To assess whether the results generalize to other study programs, data from a law bachelor program were analysed. A few differences exist in the analyses as the dependent variable, grades on a sequel second-year course, is treated as ordinal here as rounded grades are used in the law program ranging from 3 to 10 and class dependent variances were not included. If the dependent variable would be considered continuous in this case, the resulting classes would be focused too much on these eight levels and not result in relevant and insightful latent classes. Table 5 shows the validation information criteria and classification errors for the different latent class models.

The BIC, AIC, and AIC(3) values in Table 5 are all lowest for the two-class model indicating that this model fits the data best. Tests to assess the influence of the covariates on the latent classes showed that the yearly average and the variation in first-year grades had a significant influence on the latent classes, Wald statistic (1) $=$ 14.56, $p < .001$ and Wald statistic (1) $=$ 8.76, $p = .003$, respectively. As shown in Table 6, students belonging to the first class had lower average grades (6.57) than students in the second class (average of 7.5). Also, average variation in first-year grades was higher in the second class (0.87) than in the first class (0.78). Furthermore, the grade on a precursor course was a statistically significant predictor of grades on a sequel course, Wald statistic (2) $=$ 20.14, $p < .001$. Differences in the parameters across the two latent classes, however, were not statistically significant, Wald statistic (1) $=$ 1.62, $p = .200$.

As the three-class model fitted best in the psychology data and it had the second-best fit here, this model is evaluated as well. In this three-class model, only the first-year average was significantly different across the latent classes: Wald statistic (2) $=$ 20.56, $p < .001$. Furthermore, the relation between the first-year precursor course and second-year sequel course grades was significant: Wald statistic (3) $=$ 17.29, $p < .001$. The differences in the positive relation across the three classes were not statistically significant: Wald statistic (2) $=$ 3.82, $p = .150$. These results are similar to those found in the three-class model of the psychology data.

Table 5. Validation information criteria and classification errors for different LC models for law.

Mod	LL^a	BIC	AIC	AIC(3)	Number of parameters	Proportion of classification errors
1- class	−1321.90	2698.47	2659.81	2667.81	8	0
2- class	−1268.43	2680.35	2578.86	2599.86	21	0.11
3- class	−1269.22	2770.77	2606.44	2640.44	34	0.23
4- class	−1270.50	2862.15	2635.00	2682.00	47	0.17
5- class	−1268.05	2946.08	2656.10	2716.10	60	0.23
6- class	−1272.71	3044.23	2691.42	2764.42	73	0.19

[a]*LL* = Log-Likelihood. Note that the Log-Likelihood slightly increases at the 4- and 6-class model, this is possible because of the holdout validation procedure used to estimate these values.

Table 6. Descriptive statistics for the two-class model for law.

Variable		Class 1	Class 2
First-year average[a]	Average	6.57	7.5
	SD	0.41	0.52
	Max	8	9
First-year standard deviation[b]	Average	0.78	0.87
	SD	0.20	0.22
	Max	1	2
Number of compensations[c]	Average	0.91	0.23
	SD	0.93	0.49
	Max	4	2
Number of retaken tests[d]	Average	0.79	0.25
	SD	0.74	0.50
First-year precursor course grade[e]	Average	6.30	7.22
	SD	0.90	1.07
	Max	9	10
Second-year sequel course grade	Average	5.78	7.32
	SD	0.84	0.76
	Min	3	6
	Max	8	10
Class size		0.69	0.31
N[f]		286	642

[a]The minimum first-year average is six for both classes.
[b]The minimum of the first-year standard deviation is 0 in both classes.
[c]The minimum number of compensations is 0 for both classes.
[d]The range of the number of retaken tests ranges between 0 and two for both classes.
[e]The minimum of the first-year precursor course grade is 5 in both classes.
[f]Sample size here is smaller as there were 192 students for who second year grade was missing.

As shown in Table 7, the three class model, in comparison to the two-class model in which the low performing class had a first-year average of 6.4, has two classes that have a first-year average around this value, one lower and just above the required average of 6.0 and one slightly higher around 6.5. Interestingly, for the lowest performing class, the average grade on the second-year sequel course is below the Dutch pass-fail cut-score of 5.5 on average, while that of the second class is just below the required average grade of 6.0. These results of the law data show that with three classes a similar pattern is observed as in the psychology data, where low performance on the precursor course relates to an even lower performance on the sequel course on average for students whose performance in the first-year was low.

Discussion

The aim of this study was to evaluate performance on a follow-up (i.e. sequel) course that builds on material from a precursor course when students were allowed to compensate courses in the first-year of their undergraduate curriculum. The best fitting latent class model for the psychology data was a three-class model in which groups of students could be distinguished in terms of their patterns in first-year averages, number of compensated courses, and number of retakes. These three latent classes distinguish students with low, moderate and high performance. For each of the classes, the patterns are as expected: the higher the first-year average, the lower the number of compensated courses or retakes. An evaluation of the relation of performance on a precursor and sequel course show differences across the classes were not statistically significant. Overall the results suggest that students whose first-year performance is low and consequently compensate and/or fail the precursor course in the first-year have a higher likelihood to perform low on the sequel course. This positive relation is similar to findings from previous studies (e.g. Arnold 2011). Interestingly, although each class included at least some students who had to compensate the first-year precursor course, only students in the low and moderate performing

Table 7. Descriptive statistics for the three-class model for law.

		Class		
Variable		1	2	3
First-year average[a]	Average	6.42	6.71	7.25
	SD	0.37	0.46	0.63
	Max	8	8	9
First-year standard deviation[b]	Average	0.88	0.70	0.87
	SD	0.18	0.17	0.21
	Max	1	1	2
Number of compensations[c]	Average	1.5	0.43	0.45
	SD	0.93	0.63	0.72
	Max	4	2	3
Number of retaken tests[d]	Average	0.68	0.86	0.39
	SD	0.70	0.78	0.60
First-year precursor course grade[e]	Average	6.04	6.54	6.96
	SD	0.95	0.84	1.12
	Max	9	9	10
Second-year sequel course grade	Average	5.2	5.86	7.26
	SD	0.61	0.61	0.70
	Min	4	3	6
	Max	6	7	10
Class size		0.25	0.35	0.40
N[f]		230	327	371

[a]The minimum first-year average was six for all classes.
[b]The minimum first-year standard deviation was 0 for all classes.
[c]The minimum number of compensations was 0 for all classes.
[d]The number of retaken tests ranges between 0 and two for each class.
[e]The minimum first-year precursor grade was five for all classes.
[f]Sample size here is smaller as there were 192 students for who second year grade was missing.

class had to compensate the second-year sequel course. To assess the generalizability of the results, analyses were replicated using data from a law program. Results showed that similar patterns emerged in both datasets.

In this study we were specifically interested in whether latent classes could be identified in terms of similar first-year performance. We did not define these groups a priori as we did not have firm expectations on the specific formation of groups of students. We expected the latent classes to distinguish in variation in grades instead of averages, as illustrated by our example of students Ann and Peter who were hypothesized to have similar average values across courses with different degrees of variation in their grades. As shown by our resulting three latent classes with low, moderate and high performing students, this specific distinction between students like Ann and Peter was not observed as both of them would belong to the low performing class of students.

In this study, performance on a sequel course was of interest because it explicitly builds on materials from a precursor course and it gives an indication of the consequences of allowing compensation in a first-year curriculum in terms of knowledge accumulation. The results show that there is at least one group, that of the high performing students, that does not seem to make use of the compensation rule or for whom compensation does not seem to hamper performance on the second-year sequel course. These results suggest that students who compensated the precursor course, yet had higher first-year performance than students in the low performance class, are able to accumulate knowledge and skills on other courses that may transfer to the sequel course, resulting in sufficient performance.

On the other hand, the results show that for students in the low performing class (about a quarter of the sample) the average performance on the second-year sequel course is a failing grade. These results suggest that allowing course compensation might result in low performance on a sequel course when students' performance in the first-year is low as well (for psychology students characterized by a low first-year average and a high number of compensated first-year courses and retakes). This seems to suggest that allowing compensation might have negative consequences for the students in the class with overall low first-year performance, such that performance on later courses is not sufficient. In practice, this group of students with low performance implies that extra attention should be given to the requirements of allowing compensation when sequel courses exist in a curriculum. Especially when the content of these courses is considered crucial in the overall requirements of a curriculum, it might be preferred not to allow compensation for these specific precursor courses such that students' performance will not be hampered on sequel courses.

In this study, differences in patterns in students' study results were explored. These patterns might be indicative of differences in study processes. As proponents of course compensation believe compensation to positively influence students' study processes, it would be interesting for future studies to focus explicitly on these study processes. A good starting point for studying these processes could be an evaluation of how students allocate their study time. Here, experienced based sampling methods (also known as ecological momentary assessment) might provide a convenient method for measuring study time allocation. Ideally, these choices in study time allocation would be evaluated across different testing programs (i.e. compensatory or conjunctive) and in response to different curriculum aspects such as assessments and number of retakes. Also, it would be interesting to expand the model fitted in this study to different study programs to assess its generalizability, as well as to future students to assess its predictive ability. Based on our findings, which are limited only to our compensatory decision rule, it seems that students whose first-year performance is low and consequently compensate and/or fail the precursor course in the first-year have a higher likelihood to perform low on the sequel course. These students might require additional attention within a compensatory decision rule in the first-year of a curriculum to prevent gaps in knowledge, late drop-out, or increased time-to-degree.

Notes on contributors

Iris E. Yocarini was a PhD student at the Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, The Netherlands at the time of this research. Currently, she is a postdoctoral student at the Leiden Institute of Advanced Computer Science at Leiden University in Leiden, the Netherlands. Her research interests include data science (classification and predictive models), psychometrics, experimental and cognitive psychology, and applied statistics.

Samantha Bouwmeester is an Associate Professor of Methodology and Statistics at the Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, The Netherlands. Her main research interests is in psychometrics, the application of latent class models, cognitive psychology, and applied statistics.

Guus Smeets is a Professor of Psychology and former director of education at the Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, The Netherlands. His main research interest lies in education and educational psychology, as well as clinical psychology.

Lidia R. Arends is a Professor of Methodology and Statistics at the Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, The Netherlands. Besides, she is biostatistician at the Department of Biostatistics, Erasmus University Medical Center, Rotterdam, The Netherlands. Her areas of interest include research methods, (logistic) regression analysis, multilevel analysis, systematic reviews, and meta-analysis.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Iris E. Yocarini http://orcid.org/0000-0002-8592-803X

References

Arnold, I. 2017. “Resitting or Compensating a Failed Examination: Does It Affect Subsequent Results?” *Assessment & Evaluation in Higher Education* 42: 1103–1117. doi:10.1080/02602938.2016.1233520.

Arnold, I. J. M. 2011. “Compensatorische toetsing en kwaliteit [Compensatory Testing and Quality]. *Tijdschrift Voor Hoger Onderwijs* 29: 31–40. Authors, (2018). [Blinded for review].

Douglas, K.M., and R. J. Mislevy. 2010. “Estimating Classification Accuracy for Complex Decision Rules Based on Multiple Scores.” *Journal of Educational and Behavioral Statistics* 35(3): 280–306. doi:10.3102/1076998609346969.

Gruijter, D. N. M. d. 2008. Al Dan Geen Compensatie in de Propedeuse [Compensation in the First Year of the Bachelor].

Naerssen, R. F. v. 1970. *Over Optimaal Studeren en Tentamens Combineren [Oratie Universiteit Van Amsterdam]*. Amsterdam: Swets & Zeitlinger

Nuffic. 2009. “Grading systems in the Netherlands, the United States, and the United Kingdom. Den Haag: Nuffic.” https://www.nuffic.nl/en/library/grading-systems-in-the-netherlands-the-united-states-and-the-united-kingdom.pdf

Pell, G., K. Boursicot, T. Roberts. 2009. “The Trouble with Resits … .” *Assessment & Evaluation in Higher Education* 34: 243–251. doi:10.1080/02602930801955994.

Rekveld, I.J., and J. Starren. 1994. “Een examenregeling zonder compensatie in het Nederlandse Hoger Onderwijs? Een vergelijking tussen compensatie en conjunctie [An Examination Rule without Compensation in Dutch Higher Education? A Comparison between Compensation and Conjunction].” *Tijdschrift Voor Hoger Onderwijs* 12: 210–219.

Smits, N., H. Kelderman, and J. B. Hoeksma. 2015. “Een vergelijking van compensatoir en conjunctief toetsen in het hoger onderwijs [A Comparison of Compensatory and Conjunctive Testing in Higher Education].” *Pedagogische Studiën* 94: 275–285.

Sneyers, E., and K. De Witte. 2018. “Interventions in Higher Education and Their Effect on Student Success: A Meta-Analysis.” *Educational Review* 70(2): 208–228. doi:10.1080/00131911.2017.1300874.

Vermunt, J. K., and J. Magidson. 2013. *LG-Syntax Users’ Guide: Manual for Latent GOLD 5.0 Syntax Module*. Belmont, MA: Statistical Innovations.

Vossensteyn, J. J., A. Kottmann, B. W. A. Jongbloed, F. Kaiser, L. Cremonini, B. Stensaker, … S. Wollscheid. 2015. *Dropout and Completion in Higher Education in Europe: Main Report*. European Union. doi:10.2766/826962.

Wedel, M., and W. A. DeSarbo. 1994. “A Review of Recent Developments in Latent Class Regression Models.” In *Advanced Methods of Marketing Research*, edited by R. P. Bagozzi, 352–388. Cambridge, MA: Blackwell.

Yocarini, I. E., S. Bouwmeester, G. Smeets, and L. R. Arends. 2018. “Systematic Comparison of Decision Accuracy of Complex Compensatory Decision Rules Combining Multiple Tests in a Higher Education Context.” *Educational Measurement Issues and Practice* 37(3): 24–39. doi:10.1111/emip.12186.