Gutenberg School of Management and Economics

& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

Rank Response Functions in an Online Learning Environment

Tim Klausmann, Valentin Wagner and Isabell Zipperle

August 2021

Discussion paper number 2111

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
https://wiwi.uni-mainz.de/

Contact Details:

Tim Klausmann (corresponding author)
Chair of Public and Behavioral Economics
Johannes Gutenberg-University Mainz
Jakob-Welder-Weg 4
55128 Mainz
Germany
tim.klausmann@uni-mainz.de

Valentin Wagner
Chair of Public and Behavioral Economics
Johannes Gutenberg-University Mainz
Jakob-Welder-Weg 4
55128 Mainz
Germany
wagnerv@uni-mainz.de

Isabell Zipperle
Chair of Public and Behavioral Economics
Johannes Gutenberg-University Mainz
Jakob-Welder-Weg 4
55128 Mainz
Germany
isabell.zipperle@uni-mainz.de

Rank Response Functions in an Online Learning Environment

Tim Klausmann[a,*], Valentin Wagner[a], Isabell Zipperle[a]

[a] *Johannes Gutenberg University of Mainz, Johann-Joachim-Becher-Weg 31, 55128 Mainz, Germany*

Abstract

We estimate rank response functions after receiving rank-order feedback in an online learning platform. We find that the shapes of the rank response functions depend on the outcome measure under consideration. For our effort measure, i.e., whether learners continue to interact with the platform, we can reject neither a linear nor a U-shaped rank response function. For our performance measure, i.e., correctly solved exercises, we find no clear pattern overall but suggestive evidence for a linearly decreasing rank response function for individuals in the lower half of the ability distribution, i.e., the lower the rank the lower the performance.

Keywords: Rank response function, rank-order feedback, education, online learning platform
JEL: I21, D83

1. Introduction

Learners often receive feedback, including relative performance feedback, at nearly every stage of the educational system. Feedback is, according to Hattie and Timperley (2007), one of the most powerful influences on learning and achievements. However, feedback can be both positive and negative for learners' academic achievements (see, e.g., Villeval, 2020; Damgaard and Nielsen, 2018). It is therefore important to enhance our understanding under which conditions and for whom feedback works positively or negatively. This implies the need to study responses to feedback along the whole ability distribution to unmask potential non-monotonic effects.

In this paper, we estimate the response function of learners receiving rank-order feedback in an online learning platform. We exploit data of 1,009 learners who were randomly added to already existing learning groups with up to ten learners of similar ability. These are *homogeneous* feedback groups. A high-ability learner could be added, e.g., to a homogeneous feedback group of learners in the first ability quintile or, e.g., a homogeneous feedback group of learners in the fifth ability quintile. Learners received instantaneous feedback about their rank in their group after solving exercises. The online platform allows us to estimate rank response functions for two measures: effort and performance. Effort is measured by learners' continuing use of the platform after learning about their rank. Performance is the number of learners' correctly solved exercises. Our findings corroborate both, a linear and a U-shaped rank response function for effort. For our performance measure, we find suggestive evidence of a linearly decreasing rank response function only for individuals in the lower half of the ability distribution. However, overall, the shape of the rank response function for performance is inconclusive.

We contribute to the literature analyzing heterogeneous responses to relative performance feedback in educational settings. Goulas and Megalokonomou (2021) find a positive linear relationship between rank feedback and subsequent performance, i.e., the better the rank feedback in tenth grade, the higher students scored in twelfth grade. Additionally, recent evidence suggests a positive linear effect of rank on academic achievements both for university students (Elsner et al., forthcoming) and primary school students (Murphy and Weinhardt, 2020), although individuals in both studies did not directly receive feedback about their rank. In contrast to these findings, Bursztyn and Jensen (2015) find a negative linear relationship between rank feedback and performance. The performance of high-achievers decreased when a performance leaderboard was introduced in high school while low performers improved slightly. Hermes et al. (2021) find that relative performance feedback positively affects low-achieving primary students while high-achieving students did not change their behavior. Outside educational settings, there is evidence that rank response functions are non-linear. In lab studies – settings excluding social interactions between participants – Gill et al. (2019) find a U-shaped rank response function whereas Hett and Schmidt (2018) find an inversely S-shaped rank response function. Dobrescu et al. (forthcoming) is closest to our paper as they also provide in-

*Corresponding author
Email addresses: tim.klausmann@uni-mainz.de (Tim Klausmann), wagnerv@uni-mainz.de (Valentin Wagner), isabell.zipperle@uni-mainz.de (Isabell Zipperle)

stantaneous relative feedback in an online course. The authors find a constant and positive effect of relative performance feedback along the ability distribution. We differ in important ways from the above mentioned literature: we estimate rank response functions in (i) an educational online setting where (ii) peer interactions are non-existent. The latter contribution induces the advantage that our estimates of rank response functions are not biased by individuals' image concerns (see, e.g., Bursztyn and Jensen, 2015). Moreover, settings without direct peer interaction might become more common in the future due to the digitalization of education and increased usage of massive open online courses in digital workplace learning.

2. Background information and data

Setting – prüfungs.tv. Our data originate from a large online learning platform: pruefungs.tv. The content of the platform provides college-level students with videos, exercises, and mock exams. As part of a larger field study (see Klausmann, 2021), the platform implemented a homogeneous rank-order performance feedback for 2,461 registered learners. These were divided into two subsamples.

In the first subsample, learners who actively used the platform in September and/or October 2020 were allocated into homogeneous feedback groups in October and/or November 2020. As the learners' final exam was scheduled at the end of November 2020, we focus on a period of high platform activity. The allocation of learners into smaller homogeneous groups followed a two-step procedure. First, we divided learners into ability quintiles based on the number of correctly solved exercises in the previous month. Second, we randomly assigned learners within the same quintile into a group. These groups consisted of five to ten learners. Groups with less than ten learners were filled up with learners of the second subsample (see below). Learners were informed that they belong to a group and that they will be ranked based on their performance in solving exercises in the next month, i.e., the more correctly solved exercises relative to other learners the higher their rank. Groups were reshuffled each month. Importantly, the identity of learners was anonymized. There was no possibility to interact with peers, e.g., via a chat. Learners received instantaneous rank-order feedback, i.e., after correctly solving exercises they could observe their rank changing. Moreover, learners were not informed about their quintile and homogeneity of their group.

In contrast to learners in the first subsample, learners in the second subsample did not interact with the platform in September and/or October 2020 (N = 1,009). These learners were randomly assigned to a quintile and within the quintile to a homogeneous feedback group in October and/or November 2020. Thus, we do not know their initial ability level at the time of randomizing them into groups. Hence, a low-ability learner was equally likely to be sorted into one of the five already existing homogeneous quintiles. That means, learners of the same ability could end up in a stronger or weaker feedback group. This random allocation of learners of the second subsample into already existing homogeneous feedback groups is key for estimating the rank response functions.

In this study we will use the data from the second subsample only, as these students were randomly allocated to a quintile independently of their ability.

Data. We use data from 1,009 learners of the second subsample. Estimating rank response functions in our setting requires to take into account that learners were free to interact with the platform at any time and received instantaneous feedback. We therefore estimate rank response functions for different cutoff days and compare learners' effort and performance before and after the respective cutoff day with the effort and performance exerted until the end of the month. We focus on two cutoff days: (i) the *median* day where the cumulative number of activities accounted for 50% of all activities on the platform within the month (in our case the 17^{th} day) and (ii) the *mode* day which is the day with the most activities in the month, here the 23^{rd} day (for a discussion see Section 4).

Identification strategy. We apply an instrumental variable approach to estimate the rank response functions. We use the exogenous variation of randomly assigning the learners of the second subsample to feedback groups to side step potential confounds from unobserved ability and serial correlation. Without this approach serial dependence in the unobserved drivers of effort will give rise to non-causal correlation between rank at the cutoff day and future effort/performance. An example is unobserved ability that influences rank and effort in all months (compare Gill et al., 2019).

In the first stage, exemplary for the linear model presented in Panel A of Table 1, we regress rank of learner i in month t on feedback group indicators (Equation 1).

$$\widehat{rank}_{i,t}^{before} = \alpha + \rho * feedbackgroup_{i,t} + \gamma_1 * ability_{i,t} + \eta_{i,t} \quad (1)$$

In the second stage, we use a linear probability model to regress outcomes $y_{i,t}$, an indicator for *effort* or cumulative *performance* after the cutoff of learner i in month t on the rank predicted in the first stage. We control for the number of learner i's correctly solved exercises up to the cutoff day (Equation 2).

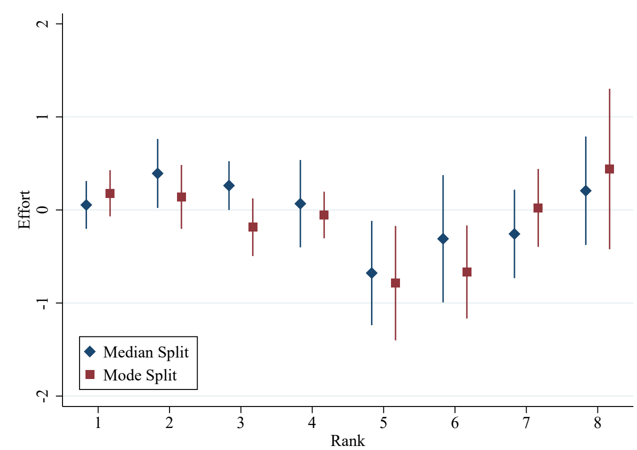
$$y_{i,t}^{after} = \beta * \widehat{rank}_{i,t}^{before} + \gamma_2 * ability_{i,t} + \epsilon_{i,t} \quad (2)$$

For the fully flexible model in Figures 1 and 2 we include indicators for each rank. Following Angrist and Pischke (2008) we satisfy the exclusion restriction as our instrument is random assignment of learners from the second subsample to feedback groups and the only information that varies between feedback groups is the rank that a learner lands in. Our instrument is also relevant, as the joint F statistic of the $feedbackgroup_{i,t}$ indicators in the first stages of the presented specifications is larger than 100.

3. Results

Figure 1 shows the fully flexible rank response functions that depicts the response to eight ranks separately (see Gill et al., 2019) for the median and the mode cutoff day. For ranks nine and ten we did not observe any participants that learned before the cutoff days. We estimate whether learners continued to use the online platform after receiving feedback on the cutoff day. Figure 1 suggests that the rank response function is broadly U-shaped which would be in line with the findings of Gill et al. (2019). Learners seem to be more likely to interact with the platform in response to rank-order feedback when they are on a high or a low rank compared to mediocre ranks.

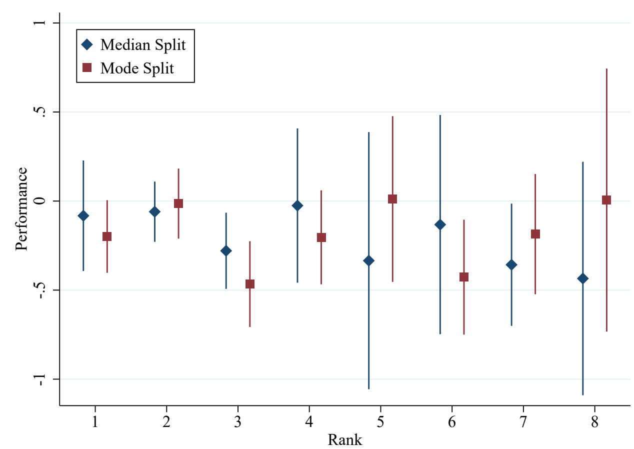
Figure 1: Fully flexible rank response function: Effort



Note: This figure presents point estimates and 95% confidence intervals of the rank response functions for effort – estimated with a linear model. Cutoff days to estimate responses to rank-order feedback are the mode day (in red; the day with the most learners active in a month) and the median day (in blue; the day where the cumulative interaction reached 50% of all interactions in a month). *Effort* indicates whether learners interacted with the platform after the cutoff day. We control for learners' ability. First place has rank 1.

Figure 2 shows the rank response functions for the median and mode cutoff day with regard to the number of correctly solved exercises (performance). Overall, we observe no conclusive shape of the rank response functions. However, Figures 2 and A.2 provide suggestive evidence that relatively to rank one and three learners ranked second and fourth respond more positively to rank-order feedback. For learners ranked in the lower half of the ability distribution (ranks five to eight), Figure 2 suggests a U-shaped rank response function when considering the mode cutoff day but a linearly decreasing rank response function when considering the median cutoff day, that is, the worse the rank the lower is the number of correctly solved exercises. Figure A.2 points to the direction that the rank response function of learners in the lower half of the ability distribution is rather linearly decreasing than U-shaped.

Figure 2: Fully flexible rank response function: Performance



Note: This figure presents point estimates and 95% confidence intervals of the rank response functions for performance. Cutoff days to estimate responses to rank-order feedback are the mode day (in red; the day with the most learners active in a month) and the median day (in blue; the day where the cumulative interaction reached 50% of all interactions in a month). *Performance* is the dependent variable measuring the number of correctly solved exercises after the cutoff day. *Performance* is normalized to 0 and first place has rank 1.

Table 1 presents the test of statistical significance for the linear (Panel A) and quadratic (Panel B) rank response function for both cutoff days and outcome variables. For effort, the coefficients for rank are negative for both cutoff days but only significant for the mode cutoff day. This suggests a linear rank response function. However, the quadratic term in column 2 of Table 1 is sizable and significant. Thus, a U-shaped rank response function with a minimum at rank 4.81 would also be feasible. For performance, we do not observe significant coefficients for both cutoff days in both specifications. Restricting our sample to learners on the lower half of the ability distribution (ranks five to eight), we find a significant linearly decreasing rank response function, the worse the rank the lower is the performance (see Table A.1 in the Online Appendix).

4. Discussion

To estimate rank response functions, we had to choose specific cutoff days to determine learners' prior and past effort and performance. We chose the median and mode day for three reasons. First, activity is high in these days (compare Figure A.1 in the Online Appendix). Second, choosing a cutoff day at the beginning of the month is not optimal since the rank at the beginning of the month with few interactions is subject to frequent change. Third, a cutoff day at the end of the month allows learners to interact with the platform only for a few days until groups are reshuffled. This would leave us with few or no observations and thus noisy estimates of effort and performance. Nevertheless, one concern might be that the shapes of the rank response functions change if we use different cutoff days. In Figures A.2 and A.3 in the Online Appendix, we show the rank response functions for each day between the 15^{th} and the 25^{th} day of the month. These figures broadly show a similar pattern as the cutoff days 'median' and 'mode'.

Table 1: Test for the quadratic rank response function

	Effort		Performance	
	Median	Mode	Median	Mode
Panel A				
Rank	-0.062	-0.087***	-0.040	-0.009
	(0.122)	(0.006)	(0.286)	(0.671)
Panel B				
Rank	0.004	-0.375***	-0.015	-0.070
	(0.976)	(0.008)	(0.908)	(0.475)
Rank^2	-0.009	0.039**	-0.003	0.008
	(0.593)	(0.028)	(0.832)	(0.495)
Joint F-test	0.185	0.004	0.484	0.774
Ability	✓	✓	✓	✓
Clus. SE	✓	✓	✓	✓
N	682	811	682	811

Note: Panel A presents estimates of the linear rank response functions and Panel B presents estimates of the quadratic rank response functions. First place has rank 1. Rank is the rank feedback received on the cutoff day and *"Sq. Rank"* is the squared rank feedback on the cutoff day. N is lower than 1,009 because not all learners interacted with the platform before the cutoff days. The cutoff day for median is day 17 and the cutoff day for mode is 23. Standard errors are clustered at the level of randomization of the instrument. 'Joint F-test' is the p-value of a joint F-test on whether Rank and Sq. Rank squared are jointly significant; p-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

5. Conclusion

Receiving feedback about how individuals stand relative to their peers is omnipresent in our daily live and its frequency is likely to increase in an ever more digitalized learning and working environment. So far, the evidence concerning rank feedback on effort and performance in education is mixed, thus increasing the need to understand heterogeneous responses along the whole ability distribution. We estimate the rank response functions of learners in a large online learning platform. Our findings indicate that the shapes of the rank response functions vary with the outcome measure under consideration. We cannot reject a U-shaped rank response function for learners' *effort*. For learners' *performance*, the rank response function seems to be constant for learners in the upper half of the ability distribution and linearly decreasing for learners in the lower half of the ability distribution. However, we do not find a conclusive overall pattern of the rank response function for the performance measure. These findings show the need for future research identifying the shape of rank response functions in educational field settings.

Acknowledgments

We would like to thank Johannes Laudenberg, Johannes Schulte, and Stephan Hansen. Moreover, we thank Daniel Schunk, Niklas Witzig, Alexander Dzionara, Patrick Schneider, and Florian Hett for helpful comments. There are no conflicts of interest; this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All errors in fact or interpretation are ours.

References

Joshua David. Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press, Princeton, NJ, 2008.

Leonardo Bursztyn and Robert Jensen. How Does Peer Pressure Affect Educational Investments? *The Quarterly Journal of Economics*, 130(3):1329–1367, 2015.

Mette Trier Damgaard and Helena Skyt Nielsen. Nudging in education. *Economics of Education Review*, 64: 313–342, 2018.

Isabella Dobrescu, Marco Faravelli, Rigissa Megalokonomou, and Alberto Motta. Relative performance feedback in education: Evidence from a randomised controlled trial. *Conditionally Accepted at The Economic Journal*, forthcoming.

Benjamin Elsner, Ingo Isphording, and Ulf Zölitz. Achievement Rank Affects Performance and Major Choices in College. *Conditionally Accepted at The Economic Journal*, forthcoming.

David Gill, Zdenka Kissová, Jaesun Lee, and Victoria Prowse. First-Place Loving and Last-Place Loathing: How Rank in the Distribution of Performance Affects Effort Provision. *Management Science*, 65(2):494–507, 2019.

Sofoklis Goulas and Rigissa Megalokonomou. Knowing who you actually are: The effect of feedback on shortand longer-term outcomes. *Journal of Economic Behavior & Organization*, 183:589–615, 2021.

John Hattie and Helen Timperley. The power of feedback. *Review of Educational Research*, 77(1):81–112, 2007.

Henning Hermes, Martin Huschens, Franz Rothlauf, and Daniel Schunk. Motivating low-achievers – relative performance feedback in primary schools. *Journal of Economic Behavior & Organization*, 187:45–59, 2021.

Florian Hett and Felix Schmidt. Pushing through or slacking off? heterogeneity in the reaction to rank feedback. 2018.

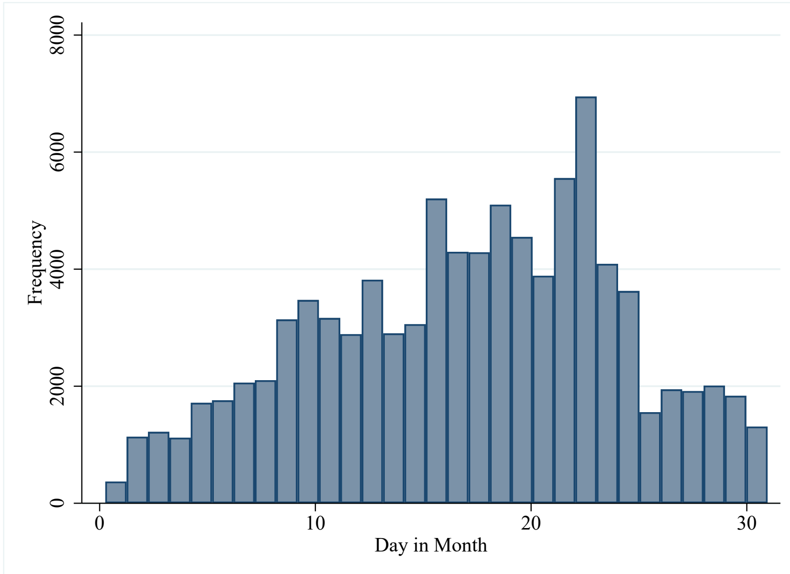
Tim Klausmann. Feedback in homogeneous ability groups: A field experiment. 2021.

Richard Murphy and Felix Weinhardt. Top of the class: The importance of ordinal rank. *The Review of Economic Studies*, 87(6):2777–2826, 2020.

Marie Claire Villeval. *Performance Feedback and Peer Effects*, pages 1–38. Springer International Publishing, Cham, 2020. ISBN 978-3-319-57365-6.

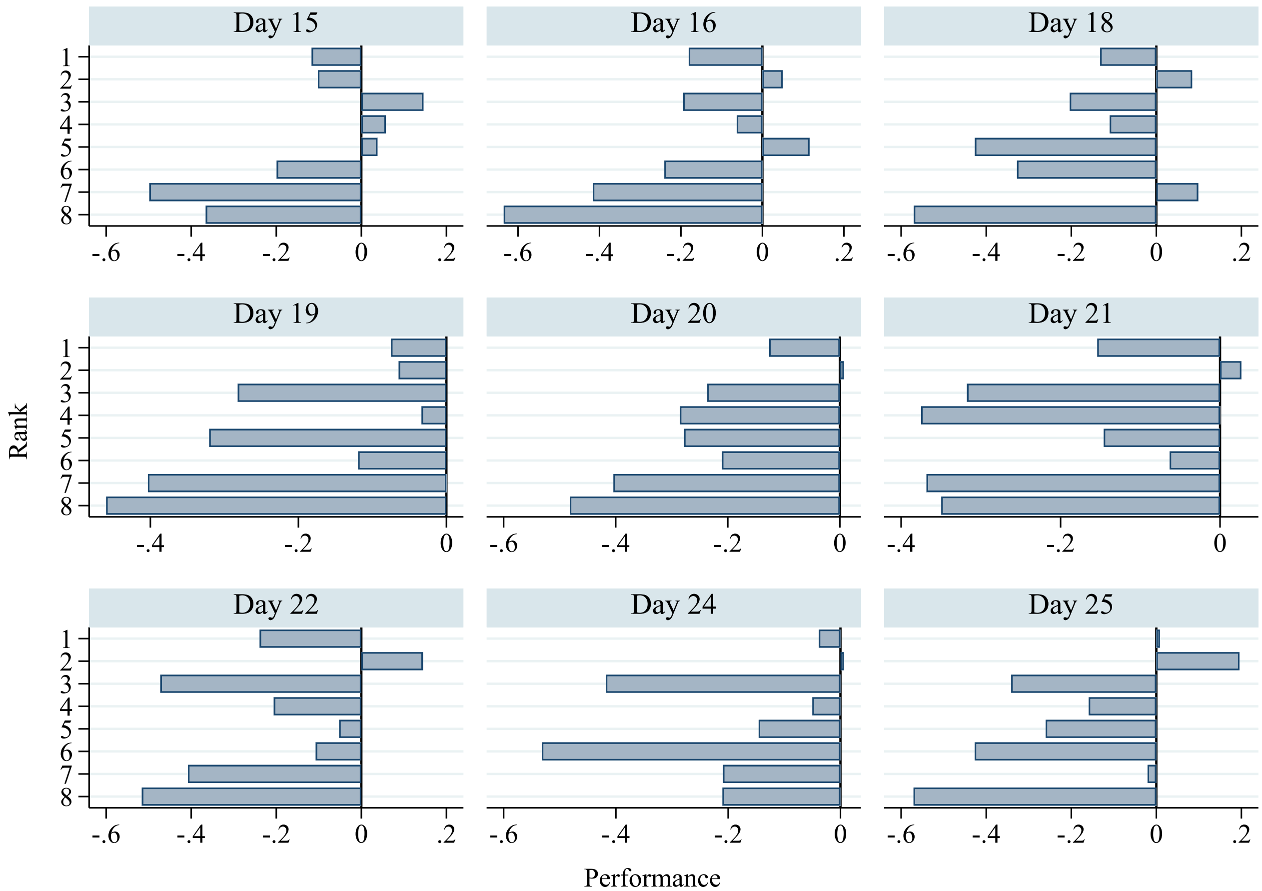
Online Appendix

Figure A.1: Activity over the course of the month.



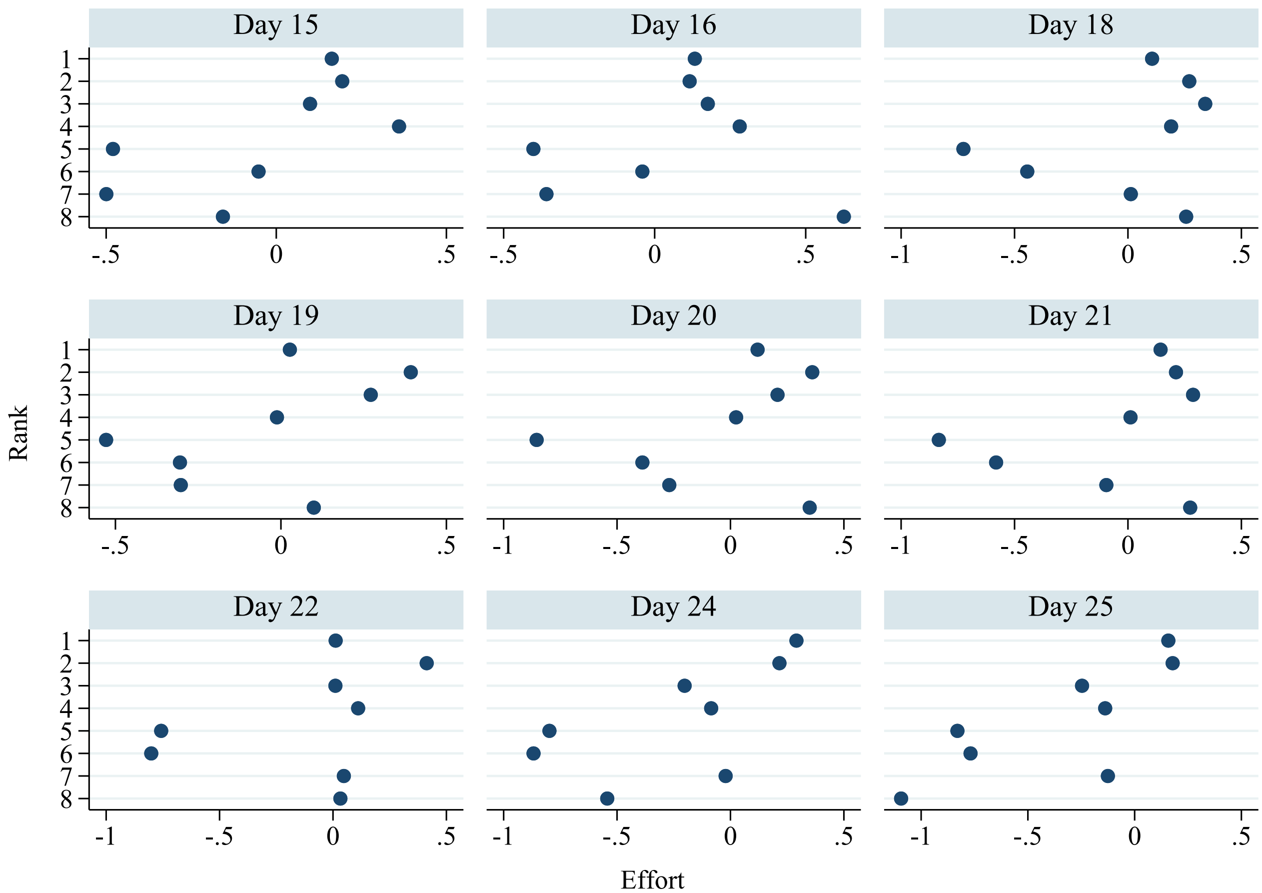
Note: This figure shows the average number of learners that logged into the platform on the given day of the month of October and November.

Figure A.2: Fully flexible rank response function for days 15 to 25: Performance



Note: This figure shows the rank response function for the performance measure for each cutoff between day 15 and day 25 in the month. Performance is the number of correctly solved exercises after the cutoff day until the end of the month.

Figure A.3: Fully flexible rank response function for days 15 to 25: Effort



Note: This figure shows the rank response function for the effort measure for each cutoff between day 15 and day 25 in the month. Effort on the x-axis indicates whether learners interacted with the platform after the cutoff day.

Table A.1: Test for the linear rank response function in the first and second half of the ability distribution (Dep. Var.: Performance)

	Median		Mode	
	Rank 1-4	Rank 5-8	Rank 1-4	Rank 5-8
Rank	-0.003	-0.065*	-0.001	-0.005
	(0.960)	(0.065)	(0.989)	(0.691)
Ability	✓	✓	✓	✓
Clus. SE	✓	✓	✓	✓
Joint F-test	0.143	0.107	0.110	0.072
N	558	124	652	159

Note: This table presents estimates of the linear rank response function for learners in the upper and lower half of the rank distribution. Dependent variable is learners' performance. Standard errors are clustered at the level of randomization of the instrument. The number of observations for ranks 5 to 8 is lower because we observe fewer active learners on low ranks in the platform. Rank is the rank feedback received on the cutoff day; p-values in parentheses. * p<0.10, ** p<0.05, *** p<0.01.