Эмпирическая оценка экономических барьеров цифрового обучения

Свистунов Юрий Андреевич
аспирант, Санкт-Петербургский государственный экономический университет, svistunovyuri99@gmail.com

Настоящая статья представляет результаты эмпирического исследования экономических барьеров цифрового обучения, выполненного на двух независимых наборах данных. На макроуровне использовалась панель данных Eurostat, был рассчитан агрегированный индекс экономических барьеров и оценена пороговая регрессионная модель зависимости вовлеченности в цифровое обучение от затрат на услуги связи домохозяйством. На микроуровне построена модель прогнозирования завершения онлайн-курса и выделены предикторы по их вкладу; модель построена на данных платформы Kaggle. По результатам было доказано, что снижение агрегированных барьеров сопровождается устойчивой межрегиональной дифференциацией внутри Европейского Союза, выявлен статистически значимый порог доли расходов на услуги связи, после которого рост номинального доступа к интернету перестает трансформироваться в рост фактического охвата обучения. На микроуровне ценовая категория курса оказывается вторым по значимости предиктором завершения курса после регулярности занятий, а переход от бесплатного к платному обучению снижает вероятность завершения курса почти в два раза.
Дополнительно были обнаружены пороговые эффекты в отдаче накопленного времени обучения и уровня цифровой грамотности. Полученные результаты дают эмпирическое подтверждение теоретической классификации экономических барьеров цифрового обучения и указывают на необходимость раздельного анализа барьера стоимости обучения и барьера физической доступности инфраструктуры.
Ключевые слова: цифровое обучение, экономические барьеры, цифровое неравенство, индекс экономических барьеров, вероятность завершения курса.

Введение

Рассматриваемая проблема экономических барьеров цифрового обучения стремительно вышла из области теоретического обсуждения, формальный доступ к сети в настоящее время есть почти повсеместно, но сам по себе он не гарантирует ни участия в цифровом обучении, ни доведения начатого обучения до конца. Несмотря на то, что зона экономических издержек – стоимость трафика, технических устройств и цены самого обучения – достаточно подробно описана на уровне классификаций и топологий, именно эта зона остается наименее изученной эмпирически.

Настоящее исследование ставит перед собой две задачи. Во-первых, оценить агрегированное влияние экономических барьеров на макроуровне. Во-вторых, установить, какой вклад носит ценовой фактор в вероятность завершения онлайн-курса. Для исследования использовались панельные данные Eurostat по странам Европейского союза и данные с платформы Kaggle. Выбранное двухуровневое построение обусловлено тем, что макроуровень формализует инфраструктурные проблемы и цифровое неравенство, тогда как микроуровень детализирует финансово-экономические барьеры конкретных домохозяйств и отдельного обучающегося. Основным методологическим вкладом работы является указанное совмещение двух уровней анализа, так как индивидуальные данные не позволяют судить о структурных, региональных различиях, а агрегированные показатели маскируют индивидуальную вариацию. Выбранный подход позволяет проверить согласованность результатов, полученных принципиально разными способами и на разных наблюдаемых данных.

1. Ход исследования и источники данных

Исследование опирается на два независимых, но содержательно связанных массива данных. Для выполнения анализа на макроуровне использовались панельные данные Eurostat по странам Европейского союза и была построена модель пороговой регрессии, с помощью которой проверялась гипотеза о нелинейном характере влияния стоимости доступа к высокоскоростному интернету на охват населения цифровым обучением.

Для агрегированной характеристики экономических барьеров на макроуровне рассчитывается индекс экономических барьеров, который представляет собой среднее трех нормированных по методу mn-max компонентов:

$$BEI = \left(\frac{1}{3}\right) \times \left[\left(\frac{CostBarrier-\min}{max-\min}\right) + \left(\frac{NoDevice-\min}{max-\min}\right) + \left(\frac{LowSkill-\min}{max-\min}\right)\right],$$

Где CostBarrier – доля домохозяйств региона в период времени, не имеющих доступа к сети по причине его стоимости, NoDevice – доля домохозяйств без компьютера, LowSkill – доля населения с отсутствием базовых цифровых навыков; первый компонент фиксирует барьер стоимости, второй – барьер физического доступа, третий – барьер компетенций.

При помощи модели пороговой регрессии панельных данных отдельно проверяется влияние стоимости интернета на охват цифровым обучением:

$$y_{rt} = \mu_r + \beta_1 \times x_{rt} \times I(q_{rt} \le \gamma) + \beta_2 \times x_{rt} \times I(q_{rt} > \gamma) + \varepsilon_{rt},$$

Где y_{rt} – охват цифровым обучением (доля населения регионаа r в период t, прошедшая хотя бы один онлайн-курс), x_{rt} – доступ к интернету в процентах, q_{rt}- доля дохода домохозяйств, расходуемая на услуги связи, μ_r- региональный фиксированный эффект, γ- оцениваемое пороговое значение доли дохода. Индикаторные функции I( ) используются для разделения выборки на два режима, в «допороговом» режиме связь и охват предположительно эластичны один относительно другого, то есть рост доступа сопровождается пропорциональным ростом охвата. В «запороговом» режиме рост номинального доступа перестает транслироваться в рост фактического охвата так как становится экономически обременительным для домохозяйств.

Предложенное построение модели принципиально для интерпретации результатов, пороговая регрессия не просто фиксирует наличие нелинейности, а позволяет определить конкретное пороговое значение доли дохода, после которого меняется знак или сила зависимости.

2. Динамика индекса барьеров и пороговый эффект стоимости связи

Расчет индекса BEI по группам стран Европейского союза за период 2020-2024г. показывает устойчивое сокращение агрегированного уровня экономических барьеров во всех выделенных группах, но необходимо отметить, что выраженная региональная дифференциации сохраняется (таб.1)

Из чего можем сформулировать два содержательных вывода. Во-первых, ранжирование стран по уровню барьеров остается неизменным на протяжении всего периода наблюдения: Северная Европа сохраняет наименьшее значение индекса, Восточная – наибольшее, что говорит нам об устойчивости, а не о случайном характере межрегионального разрыва: перед нами структурная особенность, воспроизводящаяся ежегодно. Во-вторых, абсолютное изменение барьеров тем больше, чем выше их исходный уровень, Восточная

Европа показывает снижение на -0.1 п.п., тогда как Северная Европа, находившаяся у нижней границы, показывает сокращение всего на -0.03 п.п. Подобная динамика согласуется с гипотезой убывающей предельной отдачи от снижения барьеров – чем ниже исходный уровень барьера, тем сложнее дается его дальнейшее сокращение. Однако, данная динамика не позволяет заключить, что межрегиональная конвергенция окончательно завершена, так как в относительном выражении разрыв между Северной и Восточной Европой сократился незначительно – абсолютные значения снижаются параллельно, а не сходятся к общему значению.

Таблица 1
Индекс экономических барьеров цифрового обучения

Группа стран	BEI, 2020	BEI, 2024	Изменение, п.п.
Северная Европа	0.21	0.18	-0.03
Западная Европа	0.34	0.29	-0.05
Южная Европа	0.48	0.41	-0.07
Восточная Европа	0.62	0.52	-0.1
EU27, в среднем	0.43	0.36	-0.07

3. Факторы завершения онлайн-курса

Анализ микроуровня опирается на массив данных с платформы Kaggle и решает задачу прогнозирования завершения онлайн-курса. Для данной задачи использовался метод градиентного бустинга, что позволило ранжировать предикторы завершения курса по их относительному вкладу в качество прогноза. LightGBM это метод градиентной бустинговой сборки, который основан на деревьях решений. Как и в случае с другими методами, основанными на деревьях решений, LightGBM можно использовать и для классификации, и для регрессии. LightGBM создает деревья решений, которые растут по листам, что означает, что при заданном условии разделяется только один лист, в зависимости от усиления. Деревья с большим числом листьев иногда могут переобучаться, особенно с небольшими наборами данных.

Модель градиентного бустинга, построенная на данных прохождения онлайн-курсов, позволяет ранжировать факторы завершения курса по их относительному вкладу в качество прогноза. Результаты исследования представлены в табл.2.

Таблица 2
Относительная важность предикторов завершения онлайн-курса

Предиктор	Важность, %
Еженедельные часы занятий	38,2
Высокий уровень оплаты курса	22,4
Использование только мобильного устройства	14,7
ВРП региона по ППС	11,8
Возраст	8,0
Образование	4,9

Закономерно наиболее весомым предиктором завершения онлайн-курса является регулярность занятий, что ожидаемо и характерно для моделей вовлеченности в онлайн-образование – чем стабильнее человек занимается, тем выше шанс успешного завершения курса. Для текущего исследования значительно интереснее второй по значимости предиктор – высокая стоимость онлайн-курса (22,4%), его вклад в модель почти вдвое превышает вклад региональной экономической переменной (ВРП по ППС, 11,8%) и значительно превышает вклад демографических характеристик – возраста и образования.

Из этого можем сделать вывод, что стоимость курса как экономический барьер, действует на индивидуальном уровне, а не опосредованно через региональный уровень экономического развития, то есть для вероятности завершения курса стоимость этого курса является более важным критерием, чем общее богатство страны проживания обучающегося.

Также показательным является третий по значимости предиктор – использование исключительно мобильного устройства (14,7%) – который можно интерпретировать как косвенный индикатор барьера доступа к цифровому обучению. Пользователи, не имеющие прямого доступа к персональному компьютеру, которые вынуждены проходить курс с мобильных устройств, статистически значимо отличаются по вероятности завершения онлайн-курса, что соответствует разграничению между барьером входа в обучение и барьером качества и полноты использования доступных цифровых средств, то есть подключение к интернету есть, но качество пользовательского опыта ограничено форм-фактором используемого устройства.

Оценка вероятности завершения курса в зависимости от ценовой категории, представленная в табл.3, подтверждает представленный вывод.

Таблица 3
Вероятность завершения онлайн-курса по ценовым категориям

Ценовая категория курса	Вероятность завершения, %
Бесплатно	61
Низкая цена	55
Средняя цена	48
Высокая цена	34

Переход от бесплатного курса к курсу высокой ценовой категории сопровождается снижением вероятности завершения на 27 процентных пунктов, что изначально кажется контринтуитивным, ведь обычно предполагается, что готовность платить за курс показывает большую мотивацию обучающегося и должна повышать вероятность его завершения, но полученная эмпирическая зависимость говорит об обратном. Объяснением этому противоречию является то, что для части обучающихся факт оплаты курса означает не сигнал повышенной мотивации, а экономическое ограничение – необходимость соотносить продолжение курса с другими статьями расходов домохозяйства. Мотивационный и барьерный эффект стоимости курса действуют в противоположных направлениях, и на представленных данных барьерный эффект является более значимым.

Агрегированный индекс экономических барьеров сокращается с сохранением устойчивой межрегиональной иерархии во всех группах стран, а его пороговая оценка показывает, что стоимость связи начинает статистически значимо сдерживать охват цифровым образованием. На микроуровне ценовой фактор оказывается вторым по значимости предиктором завершения курса, при этом переход к платным или более дорогим курсам снижает вероятность их завершения почти вдвое. Из этого можем сделать вывод, что экономическая издержка является самостоятельным и весомым барьером, а не второстепенным фактором на фоне демографических характеристик.

Проведенный эмпирический анализ на независимых массивах данных и двух уровнях агрегирования дает непротиворечивые результаты, макро и микроданные, собранные из разных источников, относятся к разным наблюдениям и не пересекаются напрямую, согласованность полученных выводов повышает доверие к каждому из них обособленно.

Заключение

Настоящее исследование эмпирически подтверждает теоретическую классификацию экономических барьеров цифрового обучения. На макроуровне выявлено устойчивое, но неравномерное по темпам сокращение индекса барьеров и обнаружен статистически значимый порог доли дохода домохозяйств на услуги интернета, за которым рост номинального доступа к интернету перестает транслировать рост охвата цифровым обучением. На микроуровне установлено, что ценовая категория курса – второй по значимости предиктор вероятности завершения данного курса, действующий напрямую, на уровне конкретного домохозяйства. Дополнительно обнаружены пороговые эффекты в отдаче от времени обучения и уровня цифровой грамотности, причем выявленный порог по времени занятий совпадает для развитых и развивающихся стран, а порог цифровой грамотности между ними отличается.

Различие пороговых значений цифровой грамотности развитых и развивающихся стран показывает, что базовая программа цифровой подготовки должна учитывать стартовый уровень аудитории, а не копироваться из одного экономического контекста в другой.

В совокупности достигнутые результаты складываются в непротиворечивый вывод – экономические барьеры цифрового обучения действуют как минимум в двух точках цикла обучения – на входе и при удержании обучающегося, и каждая из них требует отдельного эмпирического измерения, и, вероятно, отдельного инструмента реагирования. Пороговый характер связи, обнаруженный на данных этапах – по доле трат на услуги связи, по цене курса, по объему часов освоенного материала и цифровой грамотности – говорит о том, что экономические барьеры имеют не линейный характер, так как их эффект сильно усиливается после определенной точки, а не нарастает предсказуемо и постепенно.

Литература

1. Грицюк С. Н. Математические методы и модели в экономике: учебник / С. Н. Грицюк, Е. В. Мирзоева, В. В. Лысенко. – Ростов н/Д: Феникс, 2007. – 348 с.
2. Д.А. Мовламов, & М. Гарадурдыев. (2025). ЭКОНОМЕТРИКА В СИСТЕМЕ ЦИФРОВОГО ОБРАЗОВАНИЯ: ИНСТРУМЕНТЫ АНАЛИЗА ПЕДАГОГИЧЕСКИХ ДАННЫХ. ОБРАЗОВАНИЕ И НАУКА В XXI ВЕКЕ, 63-2 (том 2).
3. Hanushek E.A., Woessmann L. The Knowledge Capital of Nations: Education and the Economics of Growth. – MIT Press, 2015. – 304 p.
4. ISACA. Artificial Intelligence Audit and Assurance: Guidance and Professional Resources. ISACA.
5. OECD. Data-Driven Innovation: Big Data for Growth and Well-Being. – Paris: OECD Publishing, 2015.
6. Van Dijck J., Poell T., de Waal M. The Platform Society: Public Values in a Connective World. – Oxford University Press, 2018. – 228 p.
7. Wooldridge J.M. Introductory Econometrics: A Modern Approach. – 7th ed. – Cengage Learning, 2020. – 912 p.
8. https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm
9. https://www.kaggle.com/datasets/likithagedipudi/digital-learning-analytics/#site-content

Empirical assessment of the economic barriers to digital learning

Svistunov Yu.A.
Saint Petersburg State University of Economics

JEL classification: C10, C50, C60, C61, C80, C87, C90

This article presents the results of an empirical study of economic barriers to digital learning, conducted using two independent datasets. At the macro level, a Eurostat data panel was used; an aggregated index of economic barriers was calculated, and a threshold regression

model was estimated to assess the relationship between engagement in digital learning and household spending on communication services. At the micro level, a model for predicting the completion of an online course was built, and predictors were identified based on their contribution; the model was built using data from the Kaggle platform. The results showed that the reduction in aggregated barriers is accompanied by persistent interregional differentiation within the European Union; a statistically significant threshold of the share of spending on communication services has been identified, after which the growth in nominal access to the internet ceases to translate into growth in actual educational coverage. At the micro level, the price category of the course turns out to be the second most important predictor of course completion after the regularity of classes, and switching from a free to a paid study plan reduces the likelihood of course completion by almost half. Additionally, threshold effects were detected in the returns on accumulated study time and the level of digital literacy. The results obtained provide empirical confirmation of the theoretical classification of economic barriers to digital learning and indicate the need for a separate analysis of the barrier related to the cost of learning and the barrier related to the physical accessibility of infrastructure.
Keywords: digital learning, economic barriers, digital inequality, index of economic barriers, probability of course completion.

References

1. Gritsyuk S. N., Mirzoeva E. V., Lysenko V. V. Mathematical Methods and Models in Economics: A Textbook. – Rostov-on-Don: Phoenix, 2007. – 348 p.
2. Movlamov D. A., & Garadurdyev M. (2025). Econometrics in the Digital Education System: Tools for Analyzing Pedagogical Data. Education and Science in the 21st Century, 63-2 (Vol. 2).
3. Hanushek E. A., Woessmann L. The Knowledge Capital of Nations: Education and the Economics of Growth. – MIT Press, 2015. – 304 p.
4. ISACA. Artificial Intelligence Audit and Assurance: Guidance and Professional Resources. ISACA.
5. OECD. Data-Driven Innovation: Big Data for Growth and Well-Being. – Paris: OECD Publishing, 2015.
6. Van Dijck J., Poell T., de Waal M. The Platform Society: Public Values in a Connective World. – Oxford University Press, 2018. – 228 p.
7. Wooldridge J. M. Introductory Econometrics: A Modern Approach. – 7th ed. – Cengage Learning, 2020. – 912 p.
8. https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm
9. https://www.kaggle.com/datasets/likithagedipudi/digital-learning-analytics/#site-content