A comparative analysis of Udemy's catalog, pricing strategy, and offerings across providers

Suparn Padma Patra Class Central Jaipur, Rajasthan, India suparn@classcentral.com Rui Ma Shenzhen, Guangdong, China Manoel Cortes Mendez Class Central Brussels, Belgium Heba Ledwon Berlin, Germany heba.ledwon@fu-berlin.de

Abstract—Udemy is an online learning platform with 157,000 courses, 40 million users, and 428 million course enrollments. But despite its size and reach, Udemy's catalog has not been the subject of much research. In this paper, we aim to fill this gap by presenting a comparative analysis of Udemy's catalog, focusing on the provider's pricing strategy across different regions (USA, India, and France), and comparing its course offerings with those of other providers (Coursera, edX, and FutureLearn). To achieve this, we collected Udemy's course data and conducted a descriptive analysis of the provider's catalog. We then focused on Udemy's pricing model, comparing free and paid course offerings. Finally, we conducted a macro analysis of some of the features of the mentioned offerings across multiple providers and regions of the world. Among others, we found that Udemy's top-20% courses by enrollment account for about 90% of the platform's total enrollments. Udemy's free courses have a median enrollment of about 20 times that of paid courses. Udemy's list prices vary significantly across regions of the world. And while about 80% of Coursera, edX, and FutureLearn courses are free or free-to-audit, just under 10% of Udemy courses are free.

Index Terms—Udemy, Online course catalog, Comparative analysis, Pricing strategy, MOOC providers.

I. Introduction

In 2021, Udemy has over 157,000 courses in 70 languages, 40 million users, and 428 million enrollments. Despite its growing popularity among professionals and instructors, Udemy has not been the focus of much research [1]. A reason could be Udemy's different business model. Unlike Coursera, edX, and FutureLearn, which evolved from and mainly partner with academic institutions for course production, Udemy focuses on courses for jobrelated skills offered by individual instructors without the need to be affiliated with an academic institution [2].

Udemy is a peer-to-peer skill-sharing marketplace [3]. Any individual can create and offer a course on Udemy, at no charge. When a course instructor charges students for enrolling in their course, the revenue is split between Udemy and the instructor in proportion to the role

each played in acquiring the user — for instance, if the user found the course through Udemy's advertising, the provider's share will be larger. This way, the instructor can earn between 25% and 97% of the course revenue [4]. If the instructor chooses to offer the course for free, neither they nor Udemy earns revenue from the course.

Being a skill-sharing peer-to-peer marketplace, how does Udemy's business model affect its course catalog, pricing strategy, and offerings across regions? And how does this compare with other online course providers? To find answers to these questions, we present a comparative analysis of Udemy's catalog, focusing on the main features of their free and paid offerings and the provider's pricing strategy across different regions (USA, India, and France). To put our findings in context, we compare Udemy's course offerings with those of other providers (Coursera, edX, and FutureLearn).

II. METHODOLOGY

A. Data Collection

We developed a web scraper for collecting Udemy's entire course data. The scraper aggregates both data that is visible on Udemy's individual course pages and data that is only made available through Udemy's underlying API. Some of the features we collected for each course include the course title, instructor name, course level, price (if any), primary and secondary categories (subjects), course length, total enrollments, and course ratings.

We ran the scraper for the first time on January 3, 2021, collecting data for over 157,000 courses. The data collected on that day are the ones we used throughout this analysis. We ran the scraper a second time on February 2, 2021. This time, we collected course data across multiple regions of the world. More precisely, we collected Udemy's price data in three countries: the USA, India, and France. These data were used as the basis of a comparison of Udemy's global pricing strategies.

Finally, over the course of January 2021, using similar methods, we collected course data for three leading course providers: Coursera, edX, and FutureLearn. We used these

data to conduct a comparative analysis of these providers' catalogs and Udemy's catalog.

B. Data Analysis

Our exploration and analysis of the data were articulated around four axes: First, we presented an overview of Udemy's catalogue in terms of key figures. This involved aggregating courses according to individual features and highlighting potential relations between course features. For instance, we computed the average course enrollment across Udemy's different course levels. Second, we focused on Udemy's pricing strategy. In particular, we were interested in understanding the differences and similarities between Udemy's free and paid offerings. For instance, we computed the proportion of Udemy's students that were taking free courses rather than paid ones. Third, we also compared Udemy's course prices across several regions of the world.

The goal here was to understand Udemy's pricing strategy on an international scale. More specifically, we analyzed whether there were substantial price differences across three regions of the world, namely the USA, India, and France. Fourth, we expanded our analysis to other course providers. More specifically, we conducted a comparative analysis of Udemy's course catalog against those of three leading MOOC providers: Coursera, edX, and FutureLearn. We contrasted high-level features such as the total number of courses offered by each provider and the proportion of courses that are free.

III. FINDINGS

A. Catalog Overview

1) Courses: From Table I, We can see that Udemy offers 157,217 courses in total. Of those, at the time of data collection, 157,208 were marked as live (available to learners) while 9 were marked as either needing fixes or in review. Hence, for all practical purposes, virtually all courses in the dataset were readily available to learners. Therefore, this distinction will not be mentioned in the rest of this paper. About 90% of Udemy's courses are paid — only about 10% of them are free.

TABLE I Total Courses

Total number of courses	157,217
Total number of live courses	157,208
Free courses	14,900 (9.48%)
Paid courses	142,317 (90.52%)

2) Enrollments: The mean and median course enrollment shows in Table II differ significantly, suggesting a skewed distribution of course enrollments. Sure enough, when sorting courses from highest to lowest enrollment, we can see that a couple of courses have over 1 million enrollments. From there, enrollments rapidly taper off.

TABLE II COURSE ENROLLMENTS

Total enrollments	428,756,996
Mean course enrollment	2,727
Median course enrollment	188

At the other end of the spectrum, we found that 7074 Udemy courses — that is, 4.5% of the providers catalog — have no enrollments at all.

- 3) Level: Udemy offers 56,724 beginner-level courses, 20,224 intermediate-level courses and 3,331 expert-level courses. Note that an additional 76,894 Udemy courses belong to the category all levels, meaning that the instructor considers the course relevant for learners of any level.
- 4) Language: In total, Udemy offers courses in 70 different languages. In addition, courses may also include subtitles in multiple languages. Table III shows the top languages on the platform.

TABLE III
TOP LANGUAGES ON THE PLATFORM

Language	Percentage of all courses
English	63.13%
Portuguese	9.11%
Spanish	8.00%
Turkish	3.35%
German	2.89%
Japanese	2.40%
French	2.39%
Arabic	1.80%

5) Prices: The price distribution shows in Figure 1 is trimodal, with instructors favouring three prices: the minimum (20 USD), the maximum (200 USD), and around 100 USD.

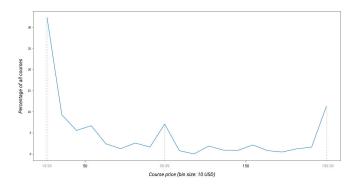


Fig. 1. Course price distribution

6) Instructors: Note that not all instructors are individuals. Many of Udemy's most prolific "instructors" are, in fact, companies that employ multiple course leads. Table IV shows that percentage of instructors with 1, 2, and 3 or more number of courses.

TABLE IV

Total number of instructors	59,032
Instructors with 1 course	61.67%
Instructors with 2 courses	16.03%
Instructors with 3 courses	22.30%
or more	22.3070

B. Comparing: Free vs Paid

In total, Udemy offers over 157,000 courses. Of those courses, close to 15,000 or 10% are free, while the rest are paid. Prices for individual Udemy courses range from \$19.99 to \$199.99. Note that our pricing data reflects list prices — that is, it does not take into account possible discounts or courses that are free due to a limited-time offer.

1) Concentrated Enrollments: When we sort Udemy's courses from lowest to highest enrollment, we notice that Udemy's enrollment distribution is skewed to the right as shown in the Figure 2, with enrollments being highly concentrated in a subset of top-performing courses. In fact, the platform's most popular courses account for the vast majority of the platform's total enrollments.

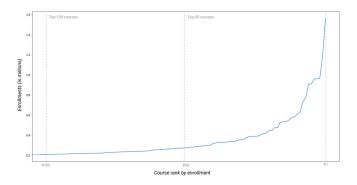


Fig. 2. Top courses by enrollment

More specifically, Udemy's top-20% courses by enrollment account for over 90% of all the enrollments on the platform. Interestingly, this seems to satisfy, and even go beyond, the Pareto 80/20 Principle. According to this principle, 80% of the impact can be attributed to just 20% of the causes. In Udemy's case, over 90% of all the enrollments can be attributed to just 20% of the courses on the platform. This unevenness carries forward to other course features, such as the number of reviews: the top-20% most reviewed courses on Udemy account for 93% of all the reviews on the platform.

2) The Rich Get Richer: This high concentration of enrollments in a small subset of courses can partly be explained by how Udemy designs its website and markets its courses. On the provider's homepage, the platform's most popular courses are featured prominently. In other

words, the marketing design creates a disproportion for the benefit of superstar courses [5].

Dedicated labels such as "Bestseller" and "Highest rated" are used to denote top-performing courses. When visiting Udemy's pages for specific categories of courses, such as web development or entrepreneurship, a carousel at the top of the page also features courses with high enrollments. Not a single one of Udemy's 15,000 free courses are shown — only paid courses are featured.

Udemy's platform seems to be designed to push users toward its already top-performing courses. As a result, courses with high enrollments tend to enjoy a steady influx of new learners, while courses with few enrollments tend to remain out of the limelight. This contributes to the uneven distribution of enrollments which we have discussed earlier. But it also leads to a "rich get richer" effect, whereby new users are not just nudged toward the platform's top-performing courses, but more specifically, toward the platform's top-performing paid courses.

Indeed, despite having a catalog with close to 10% of free courses, Udemy currently only features paid courses on its homepage. By default, Udemy's category pages and search results also show only paid courses. In fact, from Udemy's homepage, it takes at least three clicks to be able to find some of the free courses on the platform: it requires clicking on a course category, clicking on the Price filter on the left sidebar, and finally, selecting the Free checkbox. By directing learners toward its paid content, and more specifically, toward the platform's bestsellers, Udemy increases its revenue share and that of the instructors of these courses. In other words, the rich get richer.

But when it comes to free courses, Udemy has less incentive to direct learners to specific courses. Instead, other features, such as the course rating can weigh more heavily. As a result, enrollments in free courses are spread more evenly — the distribution isn't as skewed: while the top-20% paid courses by enrollment account for over 90% of all the paid enrollments, the top-20% free courses by enrollment account just for about 75% of all the free enrollments.

3) Free is Still King: In total, Udemy's course catalog has generated over 428 million enrollments. Of these enrollments, 123 million, or nearly 29% are in free courses. So, despite representing less than 10% of the courses offered by Udemy, and despite Udemy's best efforts to push learners toward its paid offerings, free courses account for over a quarter of all the enrollments on the platform. This difference is also reflected in the median enrollment: free Udemy courses have a median enrollment 20 times higher than the platform's paid courses as shown in Table V.

In addition to the inherent attractiveness of free courses, this phenomenon can partly be explained by Udemy having different strategies on and off of its platform. While on Udemy's platform the emphasis is squarely on the provider's paid courses, searching Udemy on Google returns more results related to the platform's free con-

TABLE V
ENROLLMENTS IN FREE VS PAID COURSES

	Free courses	Paid courses
Mean enrollment	8,285	2,147
Median enrollment	2,766	132

tent. This is a common SEO strategy, whereby free or discounted content can serve as a gateway toward paid offerings.

For instance, during the COVID-19 pandemic, Udemy created a resource center that lists all the free courses on the platform. The resource center is not readily accessible from Udemy's homepage. But it is one of the top results when searching "Udemy" on Google. Note, however, that the resource center currently only lists 589 free courses, while our analysis of Udemy's catalog shows that the provider actually has close to 15,000 free courses. So even on its free resources page, Udemy's does not showcase the entirety of its free offering.

4) Qualitative Differences: While the disparity in enrollments can partly be attributed to Udemy's content strategy, this only represents half of the equation. The other half are Udemy's courses themselves. Our analysis shows that at a high-level, there are noticeable differences between Udemy's free courses and paid ones. Let us discuss some of these differences.

Udemy allows its instructors to label their courses according to their level. Courses can fall in one of four categories: beginner, intermediate, expert, or all levels. Among Udemy's free courses, beginner-level courses are overrepresented. They account for 52% of Udemy's free offering, while they only represent 34% of Udemy's paid offering as shown in Table VI.

TABLE VI Course level in free VS Paid Courses

	Free courses	Paid courses
Beginner	52%	34%
Intermediate	9%	13%
\mathbf{Expert}	1%	2%
All	38%	50%

By contrast, Udemy's intermediate-level and expert-level courses show the opposite characteristic: Udemy's advanced-level courses represent a larger portion of the provider's paid offering (about 15% of all paid courses), while free intermediate-level and free expert-level courses are less common (about 10% of all free courses). This is not particularly surprising. Presumably, intermediate-level and expert-level courses require more expertise to produce, and experts may be more keen on being compensated for their experience and effort. While Udemy's website does have instructor pages, analyzing instructors' background goes beyond the purpose of this paper and is left for future research.

However, some of the measurable features of the courses themselves further speak to the qualitative differences between free and paid offerings. One of these differences is course length. From Table VII, we can see that on average, paid courses are twice as long as free ones. This also meshes with the previous observation regarding course level: more advanced courses tend to require more intricate explanations, and therefore, may tend to be longer, further justifying the attached price tag.

TABLE VII
COURSE LENGTH IN FREE VS PAID COURSES

	Free courses	Paid courses
Mean length (min)	108	258
Median length	79	137
(min)	12	137

But while we may take notice of these differences, we cannot infer from them that Udemy's paid courses are better than their free ones. In fact, the data we have compiled tend to support the opposite. Indeed, Udemy allows learners to review their courses and rate them from 1 to 5. And Udemy's free courses are, overall, rated higher than their paid courses: the median rating of free courses is 4.3, while the median rating of paid courses is 4.2 as shown in Table VIII. Interestingly, displeased learners that paid for a course may be more likely to rate it than pleased learners that took one for free, a phenomenon also noticeable in other industries [6]

TABLE VIII
RATING AND REVIEWS IN FREE VS PAID COURSES

	Free courses	Paid courses
Mean rating	4.0	3.6
Median rating	4.3	4.2
Mean number of reviews	361	187
Median number of reviews	55	12

5) Non-English Favor Free: Udemy offers courses in 70 languages. Of Udemy's 157,000 courses, over 63% are offered in English. Portuguese represents over 9% of the courses on the platform, and Spanish represents just under 8%. All the other languages represent each less than 4% of the platform's courses. One interesting aspect to note is that the proportion of free and paid courses varies depending on the language as shown in Table IX. More specifically, Udemy's free courses in English represent 8.6% of the providers total number of courses offered in English. By contrast, Udemy's free courses in Portuguese represent 11.1% of the provider's courses in Spanish represent 10.4% of the provider's courses in Spanish represent 10.4% of the provider's courses in Spanish.

This contrast is maintained when computing the same figures for all the other languages represented on Udemy's platform combined. Udemy offers 30,797 courses in languages other than English, Portuguese, or Spanish. Of

TABLE IX
COURSE LANGUAGE IN FREE VS PAID COURSES

Course language	Free courses (% of total)	Paid courses
English	8,578 (8.64%)	90,667
Portuguese	1,588 (11.08%)	12,748
Spanish	1,301 (10.35%)	11,269
Other	3,433 (11.15%)	27,364

those courses, 3,433 or 11.2% are free. In short, for every free course in English, Udemy offers about 11 paid ones, but for every free course in a language other than English, Udemy offers just about 8 paid ones. Free content is slightly more represented in non-English courses.

6) Companies as Instructors: If nearly 10% of Udemy's courses are free, only 8.4% of Udemy's instructors are dedicated to producing only free courses. In total, Udemy has over 59,000 instructors. The vast majority of them, over 83%, produce only paid courses. This leaves 8.4% of instructors who produce both free and paid offerings—with short introductory free courses often serving as a gateway toward longer and advanced-level paid courses as shown in Table X.

	Count	Percentage of total instructors
Instructors with only free courses	4,997	8.4%
Instructors with both free and paid courses	4,988	8.4%
Instructors with only paid courses	49,590	83.2%

C. Global Pricing Strategy

Udemy offers its courses internationally, and depending on the country from which you connect to the platform, you will see prices in different currencies. In total, Udemy supports 20 currencies, including the United States dollar (USD) and the euro (EUR), but also the Brazilian real (BRL), the Indian rupee (INR), and the South Korean won (KRW).

This raises a question: are Udemy's course prices equivalent in different regions of the world, or do they adjust their pricing strategy on a per-region basis — if so, to what extent? To address this question, we collected Udemy price data across multiple regions of the world. Udemy does not allow the user to select their location. Instead, they automatically detect the user's location and adjust the currency and prices accordingly. So in order to collect the price data in different regions, we had to wire our requests through multiple servers worldwide that were located in our regions of interest.

We decided to collect price data in three countries: the United States, India, and France, as a proxy to cover three of Udemy's larger areas of activity (North America, South Asia, and Europe) and three dominant currencies (USD, INR, and EUR). In future research, we may expand our data collection to additional regions and currencies.

1) USA vs India vs France: Table XI, shows that the Udemy does, in fact, have different pricing strategies for different regions. On average, courses in France and, by extension, in Europe cost 73 EUR, which at the time of data collection was equivalent to 88 USD. In the USA, however, courses only cost 73 USD on average, making Udemy's courses 17% more affordable in the USA than in Europe.

TABLE XI
COURSE PRICES ACROSS REGIONS

	France (EUR)	USA (USD)	India (INR)
Average course price (in local currency)	73	73	4682
Average course price (in USD)	88	73	64

When comparing these prices with those in India, we also find a significant difference. As we said, on average, courses cost 73 USD in the USA. In India, however, they cost on average 4682 INR, which at the time of data collection was equivalent to 64 USD. As a result, Udemy's courses in India are 12.3% more affordable than in the USA and 27.3% more affordable than in Europe. This can be seen as an example of the Penn Effect [7].

It's important to remember that this analysis focuses solely on list prices. The deltas observed between regions apply across the board to Udemy's entire catalog. They do not take into account potential discounts which are offered by individual instructors for specific courses, nor limited-time promotions by Udemy themselves.

D. Comparison with MOOC Providers

The same way that we compared Udemy's course prices across different regions, we decided to compare Udemy with other leading online course providers worldwide. More specifically, we compared Udemy's catalog with the catalogs of Coursera, edX, and FutureLearn, the three dominant MOOC platforms. Once again, the comparison is based solely on list prices and ignores courses that for a limited-time were free due to a promotion. This might have been in effect at the time of data collection, both on Udemy's side and the MOOC provider's side. In addition, the comparison focuses on high-level characteristics, such as the catalog size, the number of free courses, and their corresponding enrollments, rather than low-level features.

1) Catalog Size: At the time of data collection, Udemy offered over 157,000 courses. Coursera, its closest competitor in terms of number of courses, only offered 5,689. The platforms edX and FutureLearn trailed behind with 3,768 courses and 1,120 courses, respectively as shown in Table XII.

TABLE XII
CATALOG SIZE ACROSS PROVIDERS

	Udemy	Coursera	edX	Futurelearn
Number of courses	157,141	5,689	3,768	1,120
Number of enroll- ment	428,756, 996	154,318, 333	75,988, 351	26,830, 246

It is important to note that while Udemy offers more courses, at face value, these tend to be more skill-focused and shorter than the more academic and longer courses offered on Coursera, edX, and FutureLearn. For instance, while the workload of a Coursera course is typically described in terms of hours per week, spanning several weeks or even months, Udemy's median course length is just about two hours. A more detailed qualitative comparison of Udemy's courses with those of other providers might be the subject of future research.

We noticed that enrollments are highly concentrated in the platforms' top-performing courses, to varying degrees. Coursera's top-20% courses by enrollment account for over 80% of all the enrollments on the platform, while edX's top-20% courses by enrollment account for over 76% of total enrollments on the platform. Futurelearn has the least concentration of enrollments with its top-20% courses by enrollment accounting for over 72% of all the enrollments on the platform.

Despite being shorter than most courses typically offered by MOOC providers, Udemy's catalog still impresses by its sheer magnitude. By allowing anyone to create and offer a course through its platform, Udemy has found strength in numbers, allowing the online course provider to climb the ranking of the world's most visited websites.

2) Free Courses: While Udemy might dominate in terms of the number of courses it offers, these figures must be considered in light of the provider's course prices, accessibility, and the resulting enrollments of learners. Indeed, Udemy offers over 157,000 courses. But less than 10% of those courses are free.

From its inception, Udemy has been in the business of allowing individuals to create and sell courses. By contrast, providers such as Coursera, edX, and FutureLearn find their origins in the original MOOC movement, which emphasized openness and massiveness, for which the MOOC acronym stands for. Over time, however, MOOCs gained a business model, as providers moved some of the content that used to be free behind a paywall.

Hence, over time, the line between MOOC providers — such as Coursera, edX, and FutureLearn — and regular online course providers — such as Udemy — has become increasingly blurry.

This realization is compounded by the fact that some MOOC providers have taken concrete steps to emulate Udemy's business model. In late 2019, Coursera acquired Rhyme, an edtech startup that develops cloud-

based virtual environments. Since then, Coursera has used Rhyme software to launch Guided Projects, short project-based courses taught by subject matter experts recruited through Coursera's platform itself, similarly to how Udemy recruits its instructors.

TABLE XIII
FREE COURSES ACROSS PROVIDERS

	Udemy	Courser	$\mathbf{a} = \mathbf{edX}$	Futurelearn
Free courses	14,900	4,500	3,206	799
Percentage				
of free	9.48%	79.10%	85.08%	71.33%
courses				
Enrollment	123,359,	149,437,	71,018,	
in free	692	318	362	17,445,057
courses	092	310	302	

But the similarities stop here. Indeed, if only 10% of the courses offered on Udemy are free, over 70% of those offered by the three MOOC providers we looked at are available for free. More specifically, 85% of the courses on edX, 79% of the courses on Coursera, and 71% of the courses on FutureLearn are free or free to audit as shown in Table XIII. Therefore, although they have acquired a business model and have slowly been moving away from the openness of the Year of the MOOC [8], traditional MOOC providers' commitment to accessibility remains notable.

This difference, however, is less noticeable when looking at total course enrollments in free courses. Indeed, despite its emphasis on paid content, Udemy's free courses have a total of 123 million enrollments. Of the three MOOC providers included in our comparison, only Coursera does better, with 149 million enrollments in its free (or free to audit) courses. edX comes next with 71 million enrollments in free courses, and FutureLearn trails behind with 17 million enrollments.

IV. CONCLUSION

Hence, while its business model, its platform design, its pricing strategy, and its on-site marketing decisions put Udemy squarely in the category of "regular" online course providers, Udemy nonetheless reaches substantially more learners than both edX and FutureLearn through its free courses, further blurring the line between MOOC providers and other online course providers. How blurry this line gets, and whether it will eventually disappear, is a question we hope to explore in the future.

References

- [2] M. Conache, R. Dima, and A. Mutu, "A comparative analysis of mooc (massive open online course) platforms." *Informatica Economica*, vol. 20, no. 2, 2016.
- [3] K. Taeuscher, "Uncertainty kills the long tail: Demand concentration in peer-to-peer marketplaces," *Electronic Markets*, vol. 29, no. 4, pp. 649–660, 2019.

- [4] "Instructor revenue share udemy," https://support.udemy.com/hc/en-us/articles/229605008-Instructor-Revenue-Share, (Accessed on 05/30/2022).
- [5] E. Brynjolfsson, Y. Hu, and M. D. Smith, "Research commentary—long tails vs. superstars: The effect of information technology on product variety and sales concentration patterns," Information Systems Research, vol. 21, no. 4, pp. 736–747, 2010.
 [6] R. S. Winer and P. S. Fader, "Objective vs. online ratings: Are
- [6] R. S. Winer and P. S. Fader, "Objective vs. online ratings: Are low correlations unexpected and does it matter? a commentary on de langhe, fernbach, and lichtenstein," *Journal of Consumer Research*, vol. 42, no. 6, pp. 846–849, 2016.
- [7] M. P. Timmer, R. C. Feenstra, and R. Inklaar, "The next generation of the penn world table. forthcoming," American Economic Review, 2015.
- [8] L. Pappano, "The year of the mooc," The New York Times, vol. 2, no. 12, p. 2012, 2012.