Economics Bulletin

Volume 42, Issue 3

Independent vs major record labels: Do they have the same streaming power (law)?

Peter L Ormosi
University of East Anglia

Franco Mariuzzo
University of East Anglia

Abstract

Major and independent record labels compete for listener attention on streaming platforms. Given the superior bargaining position of major labels, biases in music recommender systems to favour more popular content, often coupled with ownership in the platform, an obvious question arises: do major record labels compete on a level playing field with independent labels in music streaming? In search of evidence this note looks at the distributional properties of the number of times a song is streamed on Spotify in the UK. We investigate whether there is a difference between the process that generates streaming numbers for UK-based major label artists and UK independents. We provide evidence of a difference in the power-law exponents of these two groups, and argue, using the scale-invariant feature of power-law distributions, that this may be a result of the difference in the streaming growth process, caused by major record labels' disproportionate presence in Spotify-generated editorial or algorithmic playlists.

We thank Steve Davies, Amelia Fletcher and Tom Gray for the useful discussions in drafting our paper. All remaining errors are ours.
Citation: Peter L Ormosi and Franco Mariuzzo, (2022) "Independent vs major record labels: Do they have the same streaming power (law)?", *Economics Bulletin*, Volume 42, Issue 3, pages 1510-1526
Contact: Peter L Ormosi - p.ormosi@uea.ac.uk, Franco Mariuzzo - f.mariuzzo@uea.ac.uk.
Submitted: February 01, 2022. **Published:** September 30, 2022.

1 Introduction

Over the past decade, streaming platforms have become central to how consumers access music content. These platforms host content from artists and receive both subscription and advertising revenue. Some of this revenue is retained by the platform (for most platforms this is around 30-40% of the total revenue), the largest part of the remaining money is paid to the record labels (recording and reproducing royalties), who then pay a fraction of what they receive to the recording artists artists. A smaller percentage of monies is also paid to publishers and songwriters (mechanical and performance royalties). Royalty payments to the rightsholders (record labels, recording artist, songwriters) are distributed based on a revenue share system, in which rightsholders receive revenues proportionate to the share of their streaming numbers to the total number of streams that the platform generates.

With this note, we would like to highlight the potential risk to competition in recorded music creation arising from the combination of the pro-rata royalty allocation method adopted by the streaming platforms, the role played by playlists in driving streams, and how these playlists are created. The intuition is as follows. Because royalty payments are proportionate to how much a song is streamed, and as inclusion in playlists is one of the most certain ways of increasing streaming numbers, a proportionate representation of record labels on these playlists is a core component of fair competition.

Computer scientists have long been studying the biases created by music recommendation systems, such as popularity bias, which favours the most popular songs in the creation of algorithmic playlists (Khenissi & Nasraoui 2020, Mansoury et al. 2020). In the economics literature, Bourreau & Gaudin (2022) argue that given the remuneration system, streaming platforms would be biased in favour of cheaper content. Closer to our research question, looking at a specific playlist (New Music Friday), Aguiar et al. (2021) do not find that a bias against independent labels exists. Our research aims to contribute to this question, but instead of focusing on specific editorial and algorithmic playlists and music charts, we ask if there can be a non-data-intensive way of evaluating the distribution of music streaming across the whole platform to get some information on potential biases. This would be an important question even in normal times, but at a time of a worldwide pandemic, which effectively shut down many revenue streams for artists, and left streaming as their almost exclusive source of revenue, the significance of this question is even more pronounced. Moreover, this question has long-lasting potential consequences. Although currently, consumers have low-price access to an unprecedented selection of music, the long-term damage can be more severe if the current revenue structure leads to a loss in music variety, as independent artists cannot recoup their investment because they are being foreclosed from receiving revenue from online streaming.

In this note, we look at streaming numbers of artists associated with UK record labels and describe the streaming growth process using the scale-invariant nature of the power-law distribution —a process that is frequently used to describe phenomena, where superstar effects are likely to dominate. Our main finding is that the power-law exponent of major label and independent label artist streams is significantly different. The scale-free property of the power-law would suggest that differences in scale (for example because major record labels sign bigger artists) are unlikely to cause a difference in the power-law exponent. Instead, the difference may come from a data generating process, other than scale. One explanation is that major record labels disproportionately feature on the biggest playlists or are disproportionately recommended by algorithmic playlists.

Another possibility is that something else (other than scale) makes major label songs different from independent ones. Although identification is difficult, we offer several experiments, which suggest that the former explanation is more likely.

2 Theoretical framework

The distribution of streaming numbers has a long right tail, with most artists receiving a relatively low number of streams, and a few popular ones getting enormous listener engagement. Several processes can describe these kinds of phenomena. An intuitively (and analytically) simple one is the power-law (also known as Pareto distribution, or 80-20 rule), which fits many human-made and natural phenomena, such as the frequency of words in books, the population of cities, the size of wars, the size of firms, earthquakes, or CEO salaries.[1] The power-law distribution has another feature, which makes it ideal for our research question: its main parameter is scale-invariant, implying that the distribution should not change with re-scaling the underlying measure.

In our context, let the (normalised) random variable number of streams S follow a power-law distribution, with counter-cumulative distribution function at s, $G(s)$:

$$P(S > s) \equiv G(s) = cs^{-k}, \tag{1}$$

where c is a constant scaler, and k is the power-law exponent.

Recalling that we have data on streaming numbers for major and independent artists, we posit that if major and independent artists are on a level-playing field on Spotify's platform, then the process of generating the streaming numbers should be similar for these two types of artists, and the only difference is expected to be in their scale (e.g. major record labels have artists with a larger fan base and wider appeal). If this assumption holds, then the main data generating power-law parameter, k, should not change due to a change in scale, and therefore should be similar for both major and independent artists. This is to do with the scale-free nature of the power-law distribution.

Assume that major label artists receive a constant λ times more streams than independents, then the distribution of the streams for these two labels should be proportional, i.e., if the distribution of indy artists' streams is given by Eq.(1), the counter-cumulative distribution of the larger label shall be $G(\lambda s) = \lambda^{-k}G(s)$. The shape of $G(s)$ is unchanged, except for an overall multiplicative constant. To put it differently, this means that the relative likelihood between songs with a few, and songs with a large number of streams in each label is the same, no matter what choice of few and many we make.

Gabaix (1999) shows that the power-law distribution can be the result of proportional growth over time. Let $S_{n,t}$ the normalised number of times song n is streamed in period t, with $S_t = \sum_{n=1}^{N} S_{n,t} = 1$, $\forall t$. Suppose that between period t and $t+1$ the number of streams of song n varies according to the random growth factor $\gamma_{n,t+1}$:

$$S_{n,t+1} = \gamma_{n,t+1} S_{n,t}, \tag{2}$$

with the growth factor being identically and independently distributed, with density $f(\gamma)$—at least in the upper part of the tail.[2] The equation of motion of the counter-cumulative distribution function $G_{t+1}(s)$ is

[1]For a summary of applications of the power-law in economics and finance, see: Gabaix (2016).

[2]The transformation $(\gamma_{n,t+1} - 1)$, is the growth rate of song n between t and $t+1$.

$$P(S_{n,t+1} > s) = P\left(S_{n,t} > \frac{s}{\gamma_{n,t}}\right) = E\left(G_t\left(\frac{s}{\gamma_{n,t}}\right)\right) = \int_0^\infty G_t\left(\frac{s}{\gamma}\right) f(\gamma)\, d\gamma, \tag{3}$$

and the steady state distribution of G is simply

$$G(s) = \int_0^\infty G\left(\frac{s}{\gamma}\right) f(\gamma)\, d\gamma. \tag{4}$$

The function $G(s) = cs^{-k}$ is a good candidate as solution of Eq. (4) since after cancelling out the terms $G(s)$ on both sides of the equation it yields

$$1 = \int_0^\infty \gamma^k f(\gamma)\, d\gamma. \tag{5}$$

The steady-state distribution is, at least in the upper tail, the power-law with exponent the k that solves $E\left(\gamma^k\right) = 1$.

If major and independent songs follow the same proportional growth process, then the fact that major record labels attract more popular artists should only change the scaling of the respective distributions. In our empirical work, we will focus on (a discrete version of) the continuous density of the power-law distribution $p(s) = cks^{-(1+k)}$ and will examine the difference in the power-law exponent k. As we observe only two waves of data, we will only be able to proxy for the growth factor of the number of streams through the cross-sectional variation in the time from the release of a song.

3 Comparing independent labels with majors

3.1 The data

Spotify does not provide access to artists' streaming history, and therefore, streaming data is only available if someone takes regular snapshots of streaming numbers on Spotify. With this note, our objective is to show that differences between major and independent artists exist, which cannot be explained only by differences in scale.

We looked at 65 independent UK record labels and the UK affiliates of the three major record labels.[3] For Warner Records UK, we only have a limited range of artists, therefore, we limit our attention to Sony and UMG. For each label, we only consider current British artists. There are 190 British artists with major record labels and 346 with independent labels in our sample. We make sure that only the albums released by their current label are included.

For each song, we collected the number of streams on two distinct dates (15 August 2020, and two weeks later, on 29 August 2020) using `chartmasters.org`. Figure 1 shows the density of streams for both major and independent record labels on a log-log scale. To plot our data, instead of treating each song individually, we arrange them into 100 bins (b), based on the number of streams each song received. These bins are equal width brackets, each of which represents a given number of songs (denoted by N_b) (for example the number of songs with streams between 0 and 1,000,000). This approach is similar to studies that use income brackets in modelling the distribution of income. Figure 1 plots the log number of songs that are in each stream bracket. To accentuate

[3] Data and codes are available at `https://github.com/PeterOrmosi/music_streaming`

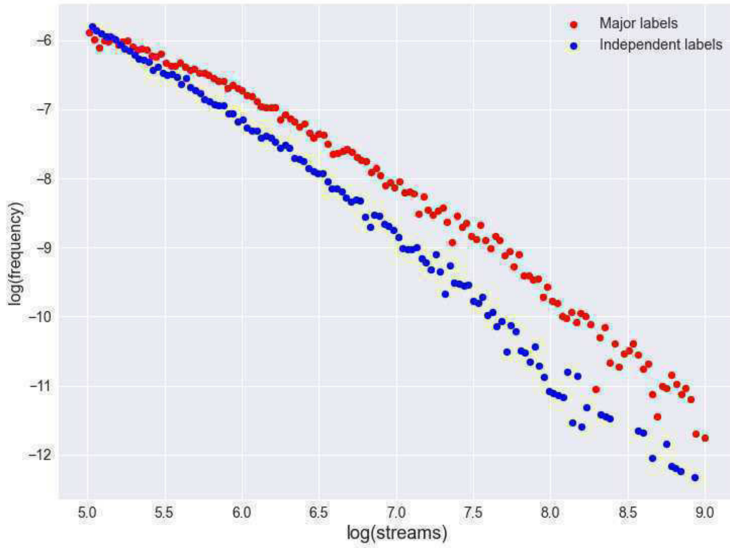

Figure 1: Empirical countercumulative distribution scatterplots for major and independent song streams

the power-law feature of our data, we focus only on the right (upper) tails of the two distributions. Figure 1 provides visual confirmation that the two distributions follow a power-law process (straight line) in the tails, and that the power law distribution is steeper for independent label artists (i.e. the relative difference between popular and less popular artists is bigger for major label artists).

3.2 Estimates

To test our main hypothesis (difference between major and indy labels) we need to estimate Eq.(1). On the left-hand side of Eq.(1) we have the counter-cumulative distribution $G(s)$, i.e. the number of songs with more streams than s. For our estimation, we replace this with the number of downloaded songs (frequency) in each stream bracket, N_b. Taking logs of each side and introducing a dummy variable for independent record labels (I^{indy}), and indicating with S_b the mid-point in the number of streams of the interval bracket b, we estimate the following log-linear model for major ($i=1$) and independent ($i=2$) record labels:

$$\ln(N_{bi}) = \beta_0 + \beta_1 I_i^{indy} + \beta_2 \ln(S_b) + \beta_3 \ln(S_b) I_i^{indy} + \varepsilon_{bi}, \tag{6}$$

where β_0 is the scaler in Eq.(1) for major label songs $(\beta_0+\beta_1)$ is the scaler for independents; β_2 is the power-law exponent for major label songs, and β_3 is the difference between the power-law exponent for independents and majors. Both β_2 and β_3 are expected to have a negative sign (negative slope of the power law curve).

Table 1 shows (for levels and growth)[4] the difference in the power-law exponent estimates

[4]For growth Eq.(6) modifies to: $\ln(N_{bi,t}) - \ln(N_{bi,t-1}) = \delta_0 + \delta_1 I_i^{indy} + \delta_2 \left(\ln(S_{b,t}) - \ln(S_{b,t-1})\right) + \delta_3 \left(\ln(S_{b,t}) - \ln(S_{b,t-1})\right) I_i^{indy} + \left(\varepsilon_{bi,t} - \varepsilon_{bi,t-1}\right)$.

between major and independent labels for British artists only —this is β_3 in Eq.(6).[5] The first two columns show the estimated difference in the exponent where we look at the total range of the distribution of stream. The third and fourth columns show the estimates where we only use the tails of the distributions with power-law characteristics. The full results are given in Table B.3 in the Appendix.

Columns 1 and 3 show the difference in the power law exponent between independent and major label artists. The negative sign indicates that the slope of the power law line is steeper (more negative) for independent labels. Because major label artists typically include artists with extremely large streaming numbers, we also looked at using a matching process to create comparable samples of indy and major label artists (Columns 2 and 4). We used a simple k-nearest-neighbour matching at the artist level (without replacement), to find the closest major label artist match to each artist in the independent sample (we matched simply based on closeness in streaming numbers). This allowed us to reduce the major sample to those artists that are most similar in their streaming numbers to independent artists. Figure A.2 in the Appendix shows how the top 1% of artists have become much more balanced as a result of the matching. The second column in Table 1 shows that the distributional difference remains (although reduced) between majors and independents.

Table 1: Power-law exponents for independent v major labels

	Total sample		Tails with power law	
	(1) indy v major	(2) indy v major (matched)	(3) indy v major	(4) indy v major (matched)
Stream (level)	-0.411***	-0.351***	-0.175**	-0.273***
	(0.053)	(0.057)	(0.068)	(0.055)
N	196	191	74	80
R2	0.95	0.948	0.985	0.982
Stream (growth)	-0.263***	-0.358***	-0.269***	-0.301***
	(0.052)	(0.053)	(0.062)	(0.062)
N	176	180	63	69
R2	0.954	0.953	0.982	0.986

The table shows the difference between the power law exponent of independent and major labels. Standard errors in parentheses.

The negative sign throughout Table 1 indicates that the slope of the power-law line is steeper (more negative) for independent labels. The obvious next question is: whether these differences are a simple reflection of the fact that major label artists are more popular, or is there something else behind our findings? Although identification is not possible from our data, the properties of the power law distribution allow us to narrow our interpretation. According to the power-law distribution, if the difference between the record labels was a simple matter of scaling (proportionately bigger or smaller), then the power-law exponent should not change.

[5] The regression estimates for all artists in our sample - and not just the British ones - are given in Tables B.1 and B.2.

To test this, we looked at several alternative scenarios. First, we compared record labels within our independent and major samples, respectively. The purpose of this exercise was to demonstrate that the difference between independent and major record labels is not simply down to differences in the size of labels. For this exercise, we created two samples. In the first one (including only independent labels) we compare the 10 largest independent labels (Cooking Vinyl, Domino Recording Company, Ninja Tune, Major League Productions (MLP), Rough Trade, Acid Jazz Records, Kscope, Moshi Moshi Records, PIAS Recordings, Smash The House) with the other 55 (smaller) independent labels. The results are reported in the first and fourth columns of Table 2. We then compared the two majors, Sony and Universal against each other (columns 2 and 5 in Table 2. The difference in the power-law exponent is not significant for either of these comparisons (neither are the scalers). This would suggest that the difference between major and independent that we found in Table 1 is not simply due to a difference in the popularity of different artists, but something else, probably to do with the data generating growth process, as posited in our theoretical framework.

As another scenario, we also compared the top independent (highest streaming independent artists) with the major artists (columns 3 and 6). This was to verify whether the most-streamed independent artists are similar to the major artists. We found that the power-law coefficients are still statistically significantly different.[6]

Table 2: Experiments with alternative scenarios

	Total sample			Tails with power law		
	(1) indy (top) vs indy (other)	(2) sony vs universal	(3) indy (top) vs major	(4) indy (top) vs indy (other)	(5) sony vs universal	(6) indy (top) vs major
Stream (level)	0.035	0.028	-0.366***	0.135	-0.017	-0.175**
	(0.047)	(0.059)	(0.055)	(0.089)	(0.071)	(0.068)
N	182	197	196	65	69	74
R2	0.971	0.912	0.943	0.983	0.978	0.985
Stream (growth)	-0.002	0.046	-0.224***	-0.066	0.046	-0.366***
	(0.047)	(0.057)	(0.054)	(0.063)	(0.057)	(0.055)
N	174	166	176	59	66	69
R2	0.971	0.93	0.948	0.98	0.93	0.986

The table shows the difference between the power law exponent of different comparison groups as shown in the heading. Standard errors in parentheses.

Finally, we tested the difference between the releases of the same artists. In one group we put those albums of currently independent artists that were released by independent labels, and in the other group, we have albums of the same artists that were released by majors (if available). The power-law exponent of streaming numbers is significantly different, with major releases having a fatter tail as shown in Table 3.[7] All of these tests would support the hypothesis that there is an

[6] Table B.4 in the Appendix shows the full regression results.

[7] Table B.5 in the Appendix shows the full regression results.

underlying difference in the data generating process beyond the fact that independent artists have a smaller listener base than those signed up with major record labels.

Table 3: Comparison of independent artists based on release label

	Full sample	Tails with power law
Stream count	-0.221***	-0.286***
	(0.052)	(0.058)
N	190	83
R2	0.962	0.982
Stream growth	-0.202***	-0.234***
	(0.058)	(0.057)
N	176	72
R2	0.959	0.986

The table shows the difference between the power law exponent of the independent and major releases of the same artists. Standard errors in parentheses.

3.3 Discussion and conclusion

We found that the relative likelihood of observing a song with many streams in comparison to a song with few streams is the same across all independent labels, and across all major labels, but it is different if one compares independent with major labels. As an alternative scenario, we show that there is no difference in the distribution if one compares the most popular independent labels with less popular independent record labels. On the other hand, the difference between major and independent exists even if one compares independent and major releases of the same artists, or the most popular independent labels with major labels. Due to the scale-invariant nature of the power-law distribution, these results suggest that there is more to the difference between independent and major labels than what would simply be dictated by differences in scales.

There are numerous reasons why we believe this finding is do to major labels' disproportionate access to these playlists. In the case of editorial playlists (playlists curated by an editor), major labels have have enormous bargaining advantage due to the sheer size of the catalogues they represent. As these editors often engage in pitching sessions with the record labels, this gives major labels the opportunity to take advantage of their bargaining position. Spotify's incentives to playlist songs from the major labels may also be influenced by their contracts with those labels. While these are confidential, Spotify states that they include minimum payment guarantees, which require it to make payments even if that label's recordings do not hit a specified level of streams. Putting more of that label's music onto playlists would clearly reduce the risk of triggering such payments.

Regarding algotorial playlists (playlists created by an algorithm), there is a rich literature studying the biases caused by autonomous recommender systems (such as the ones generating the algotorial playlists). For example, popularity bias means that the largest and most popular labels/artists are disproportionately recommended to listeners. The main point is that through feedback loops, even the smallest of these biases can enormously tilt the playing field towards major labels. This is not a speculative academic point, streaming platforms are aware of this problem (Mehrotra et al. 2018). Finally, major labels often have direct or indirect ownership in the streaming platforms. For example, Sony and Universal have direct ownership in Spotify. Warner and Sony have indirect ownership through Tencent. Whilst these are minority shares, it is not far-fetched to argue that this may distort incentives in creating a balanced/impartial playlist. Finally, some of the playlists owned and curated by the major labels (for example around 7% of the top 1000 Spotify playlists are owned by the major labels).

Due to the importance of playlists in driving streams and thereby revenue for artists, any potential advantage in being included in the most popular playlists can have important consequences for music suppliers. As such, it would be paramount in the interest of a level playing streaming field to provide greater transparency as to how streaming platforms' proprietary and algorithmic playlists are created. This improved transparency would also facilitate market entry by helping the development of a competitive and innovative market for third-party playlist creation. To support these calls for greater transparency, in our currently ongoing work programme we look in more detail at the different reasons why the competitive playing field between music companies may be tilted on streaming platforms, which includes simulations on how recommender systems affect competition, and a detailed randomised experiment to provide evidence of bias against certain types of music companies. These works will also include a much large dataset, to verify how much we can generalise from the findings in this note.

References

Aguiar, L., Waldfogel, J. & Waldfogel, S. (2021), 'Playlisting favorites: Measuring platform bias in the music industry', *International Journal of Industrial Organization* **78**, 102765.

Bourreau, M. & Gaudin, G. (2022), 'Streaming platform and strategic recommendation bias', *Journal of Economics & Management Strategy* **31**(1), 25–47.

Gabaix, X. (1999), 'Zipf's law for cities: an explanation', *The Quarterly journal of economics* **114**(3), 739–767.

Gabaix, X. (2016), 'Power laws in economics: An introduction', *Journal of Economic Perspectives* **30**(1), 185–206.

Khenissi, S. & Nasraoui, O. (2020), 'Modeling and counteracting exposure bias in recommender systems', *arXiv preprint arXiv:2001.04832* .

Mansoury, M., Abdollahpouri, H., Pechenizkiy, M., Mobasher, B. & Burke, R. (2020), Feedback loop and bias amplification in recommender systems, *in* 'Proceedings of the 29th ACM International Conference on Information & Knowledge Management', pp. 2145–2148.

Mehrotra, R., McInerney, J., Bouchard, H., Lalmas, M. & Diaz, F. (2018), Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness & satisfaction in recommendation systems, *in* 'Proceedings of the 27th acm international conference on information and knowledge management', pp. 2243–2251.

Online appendix

A Data description

Our data collection resulted in 748 artists at major record labels (Sony and UMG) and 1,105 artists at independents.[8] For each artist, we looked at their songs available on Spotify. This gave us 96,701 songs by artists at majors and 96,820 songs by artists with independents. For each song, we collected the number of streams on two distinct dates (15 August 2020, and two weeks later, on 29 August 2020). In the main part of this paper, we focus only on the 190 British artists with major record labels and 346 with independent labels. This online Appendix offers results for the total sample as well. Our headline results remain qualitatively the same.

Table A.2 summarises the main features of our data. In our sample, by 15 August 2020 major label artists received a total of 1,193 billion streams (an average of 1.66 billion streams per artist), whilst independent artists got a more modest 102 billion streams (an average of 92 million streams per artist). This means that in our sample of UK record labels, over 90% of the streams were enjoyed by major record labels. Streaming numbers increased by a total of 11.4 billion for major artists (average 15.87 million per artist) and by 897 million for independent artists (average 820 thousand per artist). Needless to say, the average figures are hugely skewed by the presence of a small number of largely successful artists (both major and independent). To demonstrate this, Figure A.1 in the Appendix shows the number of streams for the top 1% of most-streamed artists (accounting for 22% of the total number of streams), all of whom are with major record labels. UK-based major label artists had an average of one billion streams and a total of 1,973 billion streams. UK-based indy artists received a total of 42 billion streams (an average of 120 million per artist). Streaming data by the record labels in our sample are given in Table A.1.

Finally, our raw data represented all previous works of a given artist under the label that they are currently associated with. In reality, it is often the case that artists might start with an independent label and migrate to a major, or the other way around. For example, the first album of the British artist Fryars was released by the independent frYarcorp in 2008, and his second album, 'Power' in 2014 by Fiction Records, which is part of one of the majors, UMG. One argument could be that once with the majors, the major label becomes incentivised to improve all streaming revenue for their artists, irrespective of whether it is the streaming of one of their releases, or an earlier release under an independent record label. On the other hand, streaming numbers of independent label releases may remain different from major releases even after the major label signs the same artist. For this reason, we worked with two samples of data. In the first one, and the one we highlight

[8]The list of independent labels in our sample: 3 Beat Records, 4AD, ATP Recordings, Acid Jazz Records, Alcopop! Records, AudioPorn Records, Audiobulb Records, Best Before Records, Big Scary Monsters Recording Company, Bloody Chamber Music, Citinite, Convivium Records, Cooking Vinyl, Cult Records, Deltasonic, Dented Records, Dirty Hit, Domino Recording Company, Dreamboat Records, Erased Tapes Records, Fat Cat Records, Fire Records (UK), Full Time Hobby, Gringo Records, Hassle Records, Heavenly Recordings, Heist Or Hit Records, Holy Roar Records, Hospital Records, Kitchenware Records, Kscope, LAB Records, Last Night From Glasgow, Lojinx, Loose Music, LuckyMe (record label), Major League Productions (MLP), Marrakesh Records, Memphis Industries, Moshi Moshi Records, Ninja Tune, One Little Indian Records, PIAS Recordings, Peacefrog Records, Pickled Egg Records, Platform Records, Rephlex Records, Rise Above Records, Rock Action Records, Rough Trade, Smalltown America, Smash The House, Snakes & Ladders Records, Sons Ltd., Southern Records, Stolen Recordings, Tigertrap Records, Tin Angel Records, Transcend Music, Transgressive Records, Visible Noise, Wichita Recordings, XL Recordings, Xtra Mile Recordings, Young Turks.

throughout the paper we correct the labels, and in the second one we assume that all releases of an artist belong to their current label. For comparison, our total sample and the sample where the labels are corrected are both included in Table A.2 above.

To do this correction, for each album in our sample, we downloaded information on their release from `musicbrainz.org`. This included information on the label that released the given album. Where an album was released in multiple countries, the corresponding label for each of these countries is given. We also downloaded the list of labels associated with the major record labels.[9] We then searched our album release information to detect if one of these major record labels were associated with any of our albums.[10] In this matching process, we discarded all observations where we were not certain whether the album was released by an independent or a major record label. This resulted in a reduced sample of 637 major labels (176 British), and 1,096 independent label artists (344 British), including 23,663 songs by artists at majors (7,554 for British artists), and 91,086 songs by artists with independents (34,456 by British artists). These numbers show that the dominant artist trajectory goes from artists releasing their earlier albums with independents and then joining a major record label. In the empirical work that follows we only used this corrected sample.

[9] `https://en.m.wikipedia.org/wiki/List_of_Warner_Music_Group_labels`, `https://en.m.wikipedia.org/wiki/List_of_Universal_Music_Group_labels`, `https://en.m.wikipedia.org/wiki/List_of_Sony_Music_labels`

[10] As the name of the record label - as recorded on musicbrainz.com varies significantly, we used fuzzy string matching. Through this matching, we identified a similarity score and were able to set a threshold, beyond which, all matches were accurate. If at least one of the labels that released the given album/song was a major, then the album/song is classified as major.

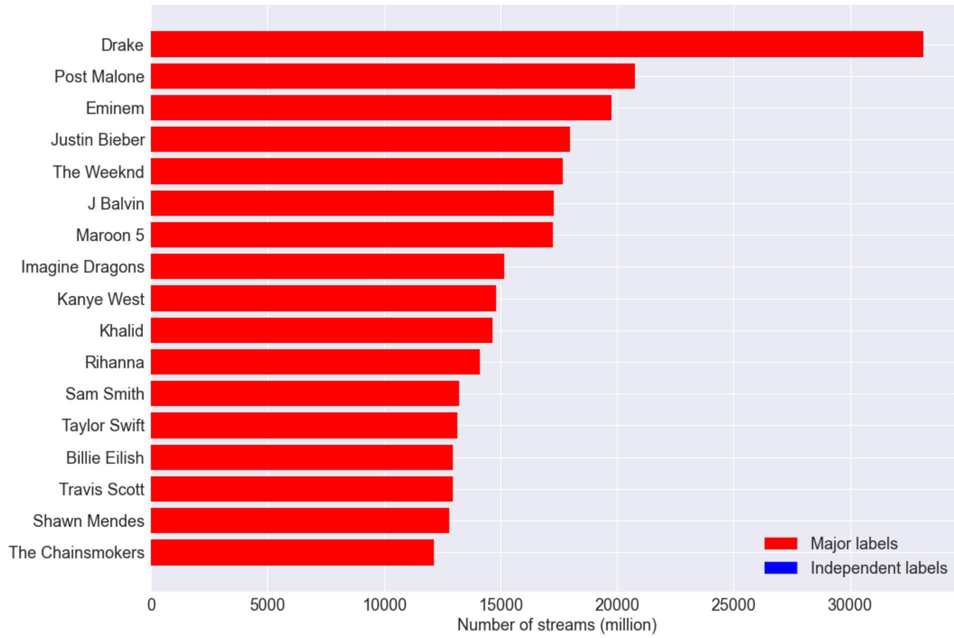

Figure A.1: Number of streams for the top 1% of most streamed artists

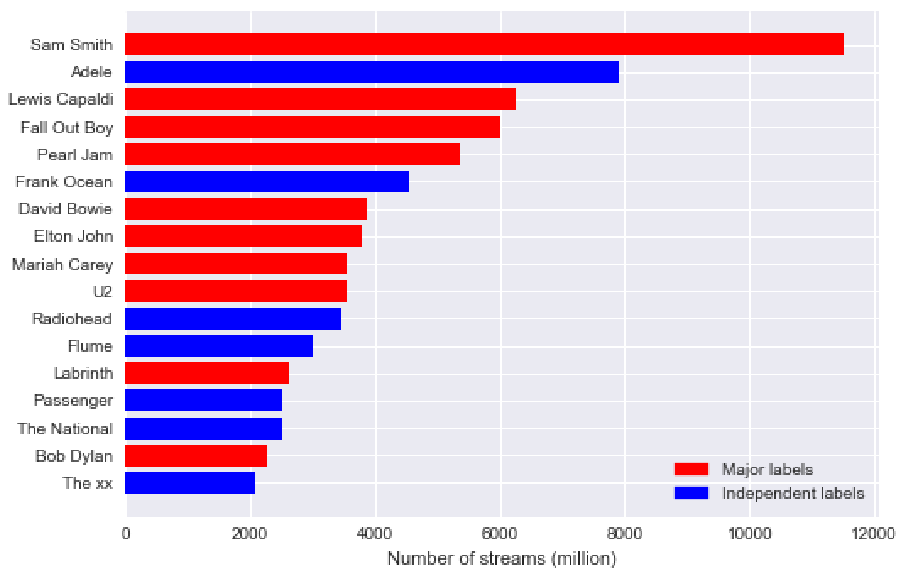

Figure A.2: Number of streams for the top 1% of most streamed artists - matched sample

Table A.1: Major and Independent labels

	Number of artists	Number of songs	Total streams (million)	Stream/song (million)
MAJOR LABELS				
Sony	346	48333	512607.732	10.606
UMG	410	48368	690058.547	14.267
Warner	26	2689	54183.552	20.150
INDEPENDENT LABELS				
3 Beat Records	24	2725	5246.439	1.925
4AD	27	2400	8069.473	3.362
ATP Recordings	12	458	133.513	0.292
Acid Jazz Records	43	4240	699.768	0.165
Alcopop! Records	19	1231	335.081	0.272
AudioPorn Records	13	639	262.072	0.410
Audiobulb Records	5	89	23.682	0.266
Best Before Records	8	266	16.783	0.063
Big Scary Monsters Recording Company	33	2501	705.742	0.282
Bloody Chamber Music	2	102	24.697	0.242
Citinite	1	61	0.301	0.005
Convivium Records	10	627	11.947	0.019
Cooking Vinyl	60	10104	10138.204	1.003
Cult Records	8	363	599.764	1.652
Deltasonic	11	492	231.741	0.471
Dented Records	3	360	61.101	0.170
Dirty Hit	17	609	1606.317	2.638
Domino Recording Company	53	6546	7002.010	1.070
Dreamboat Records	6	209	60.566	0.290
Erased Tapes Records	17	1456	1022.514	0.702
Fat Cat Records	21	1972	470.820	0.239
Fire Records (UK)	16	1846	68.445	0.037
Full Time Hobby	16	1065	184.454	0.173
Gringo Records	6	98	0.683	0.007
Hassle Records	10	878	1253.341	1.427
Heavenly Recordings	17	1492	392.797	0.263
Heist Or Hit Records	5	60	8.595	0.143
Holy Roar Records	12	668	43.044	0.064
Hospital Records	18	1989	1325.047	0.666
Kitchenware Records	4	72	0.702	0.010
Kscope	28	3703	768.952	0.208
LAB Records	14	475	397.730	0.837
Last Night From Glasgow	11	160	1.165	0.007
Lojinx	26	2564	530.730	0.207
Loose Music	17	1501	438.102	0.292
LuckyMe (record label)	18	1339	743.241	0.555
Major League Productions (MLP)	13	4143	1348.389	0.325
Marrakesh Records	3	116	2.588	0.022
Memphis Industries	34	2368	432.455	0.183
Moshi Moshi Records	49	3614	4451.477	1.232
Ninja Tune	37	4316	3995.241	0.926
One Little Indian Records	10	670	47.288	0.071
PIAS Recordings	28	3792	4894.023	1.291
Peacefrog Records	12	1269	211.573	0.167
Pickled Egg Records	16	883	170.285	0.193
Platform Records	4	84	0.729	0.009
Rephlex Records	18	1970	88.152	0.045
Rise Above Records	11	324	26.572	0.082
Rock Action Records	9	765	261.173	0.341
Rough Trade	52	4340	5398.638	1.244
Smalltown America	7	135	3.456	0.026
Smash The House	17	2846	6724.660	2.363
Snakes %26 Ladders Records	1	18	0.403	0.022
Sons Ltd.	1	59	0.225	0.004
Southern Records	27	839	141.842	0.169
Stolen Recordings	10	434	35.363	0.081
Tigertrap Records	9	252	9.085	0.036
Tin Angel Records	18	869	15.852	0.018
Transcend Music	19	757	78.051	0.103
Transgressive Records	19	1151	4621.539	4.015
Visible Noise	2	57	8.202	0.144
Wichita Recordings	20	1702	831.937	0.489
XL Recordings	23	1575	18965.786	12.042
Xtra Mile Recordings	32	1799	330.224	0.184
Young Turks	9	313	3372.808	10.776

Table A.2: Summary statistics

	Total sample	Total sample (labels corrected)	British artists	British artists (labels corrected)
Number of independent labels	65	65	58	58
Number of major artists	719	637	190	176
Number of independent artists	1101	1096	346	344
Average number of songs - major artists	127.28	37.15	141.86	42.92
Total number of songs - major artists	91514	23663	26953	7554
Average number of songs - indy artists	91.57	83.11	111.97	100.16
Total number of songs - indy artists	100814	91086	38742	34456
Average number of streams (million) - major artists	1659.80	704.85	1038.30	503.20
Total number of streams (million) - major artists	1193397.63	448987.69	197277.70	88563.46
Average number of streams (million) - indy artists	92.74	91.48	120.52	110.52
Total number of streams (million) - indy artists	102108.21	100261.24	41698.39	38019.82
Average stream growth (million) - major artists	15.87	6.71	9.45	4.61
Total stream growth (million) - major artists	11410.85	4276.61	1795.40	811.41
Average stream growth (million) - indy artists	0.82	0.80	0.98	0.90
Total stream growth (million) - indy artists	897.50	880.30	339.09	310.73

B Other figures and tables

Table B.1: power-law exponents for independent v major record labels (all artists)

	Total sample		Tails with power law	
	(1) indy v major	(2) indy v major (matched)	(3) indy v major	(4) indy v major (matched)
Stream (level)	-0.44***	-0.287***	-0.075	-0.313***
	(0.055)	(0.057)	(0.128)	(0.072)
Stream (growth)	-0.254***	-0.288***	-0.001	-0.08
	(0.055)	(0.056)	(0.125)	(0.103)

The table shows the difference between the power law exponent of independent and major labels. Standard errors in parentheses.

Table B.2: Experiments with alternative scenarios (all artists)

	Total sample			Sample with power law		
	(1) indy (top) vs indy (other)	(2) sony vs universal	(3) indy (top) vs major	(4) indy (top) vs indy (other)	sony vs universal	indy (top) vs major
Stream (level)	-0.025	0.001	-0.262***	0.084	-0.114	-0.048
	(0.05)	(0.058)	(0.061)	(0.067)	(0.132)	(0.128)
Stream (growth)	-0.01	0.044	-0.106*	-0.066	0.027	-0.262***
	(0.052)	(0.058)	(0.059)	(0.063)	(0.057)	(0.061)

The table shows the difference between the power law exponent of different comparison groups as shown in the heading. Standard errors in parentheses.

Table B.3: Full estimates for independent v major labels (British artists only)

	Total sample		tails with power law	
	major v indy	major v indy (matched)	major v indy	major v indy (matched)
Stream (level)				
scaler (major)	-1.738***	-1.976***	-4.141***	-3.69***
s.e.	(0.207)	(0.201)	(0.162)	(0.106)
scaler (indy v major)	-2.076***	-1.837***	-1.234***	-1.684***
s.e.	(0.258)	(0.253)	(0.22)	(0.182)
exponent (major)	-1.065***	-1.124***	-1.81***	-1.712***
s.e.	(0.041)	(0.047)	(0.052)	(0.032)
exponent (indy v major)	-0.411***	-0.351***	-0.175**	-0.273***
s.e.	(0.053)	(0.057)	(0.068)	(0.055)
N	196	191	74	80
R2	0.95	0.948	0.985	0.982
Stream growth				
scaler (major)	-2.37***	-1.604***	-4.032***	-3.349***
s.e.	(0.206)	(0.196)	(0.147)	(0.166)
scaler (indy v major)	-1.226***	-1.991***	-1.176***	-1.859***
s.e.	(0.263)	(0.255)	(0.19)	(0.205)
exponent (major)	-1.137***	-1.042***	-1.66***	-1.63***
s.e.	(0.039)	(0.041)	(0.049)	(0.05)
exponent (indy v major)	-0.263***	-0.358***	-0.269***	-0.3***
s.e.	(0.052)	(0.053)	(0.062)	(0.062)
N	176	180	63	69
R2	0.954	0.953	0.982	0.986

The table shows the difference between the power law exponent of independent and major labels.
Standard errors in parentheses.

Table B.4: Full estimates for experiments with alternative scenarios (British artists only)

	Total sample			Tails with power law		
	indy (top) v indy (other)	sony v umg	indy (top) v major	indy (top) v indy (other)	sony v umg	indy (top) v major
Stream (level)						
scaler (major)	-3.566***	-1.78***	-1.738***	-5.381***	-4.144***	-4.141***
s.e.	(0.187)	(0.216)	(0.207)	(0.291)	(0.169)	(0.162)
scaler (indy v major)	0.229	0.217	-1.843***	0.716**	0.139	-1.147***
s.e.	(0.23)	(0.294)	(0.268)	(0.302)	(0.225)	(0.22)
exponent (major)	-1.43***	-1.066***	-1.065***	-1.953***	-1.782***	-1.81***
s.e.	(0.037)	(0.045)	(0.041)	(0.086)	(0.052)	(0.052)
exponent (indy v major)	0.035	0.028	-0.366***	0.135	-0.017	-0.175**
s.e.	(0.047)	(0.059)	(0.055)	(0.089)	(0.071)	(0.068)
N	182	197	196	65	69	74
R2	0.971	0.912	0.943	0.983	0.978	0.985
Stream growth						
scaler (major)	-3.315***	-1.241***	-2.325***	-4.854***	-1.241***	-3.349***
s.e.	(0.172)	(0.187)	(0.207)	(0.127)	(0.187)	(0.166)
scaler (indy v major)	0.041	-0.08	-1.0***	0.043	-0.079	-1.859***
s.e.	(0.232)	(0.261)	(0.278)	(0.212)	(0.261)	(0.205)
exponent (major)	-1.348***	-1.093***	-1.126***	-1.784***	-1.093***	-1.63***
s.e.	(0.035)	(0.04)	(0.039)	(0.038)	(0.04)	(0.05)
exponent (indy v major)	-0.002	0.046	-0.224***	-0.066	0.046	-0.3***
s.e.	(0.047)	(0.057)	(0.054)	(0.063)	(0.057)	(0.062)
N	174	166	176	59	166	69
R2	0.971	0.93	0.948	0.98	0.93	0.986

The table shows the difference between the power law exponent of different comparison groups as shown in the heading.
Standard errors in parentheses.

Table B.5: Full estimates for comparison of independent artists based on release label (British only)

	Total sample	Tails with power law
Stream (level)		
scaler (major)	-2.341***	-3.643***
s.e.	(0.162)	(0.123)
scaler (indy v major)	-1.473***	-1.732***
s.e.	(0.224)	(0.193)
exponent (major)	-1.254***	-1.699***
s.e.	(0.04)	(0.037)
exponent (indy v major)	-0.221***	-0.286***
s.e.	(0.052)	(0.058)
N	190	83
R2	0.962	0.982
Stream growth		
scaler (major)	-2.258***	-3.68***
s.e.	(0.19)	(0.137)
scaler (indy v major)	-1.338***	-1.479***
s.e.	(0.251)	(0.187)
exponent (major)	-1.198***	-1.678***
s.e.	(0.047)	(0.042)
exponent (indy v major)	-0.202***	-0.234***
s.e.	(0.058)	(0.057)
N	176	72
R2	0.959	0.986

The table shows the difference between the power law exponent of the independent and major releases of the same artists. Standard errors in parentheses.