



Prepared by Will Page,  
Chief Economist PRS for Music  
and Eric Garland, CEO BigChampagne



#### About Eric Garland

Eric Garland is co-founder and Chief Executive Officer of BigChampagne Media Measurement, a privately-held technology and metrics company specialising in online media. At present, BigChampagne's available metrics include online retailers (iTunes), social networks (MySpace, YouTube), portals Yahoo and AOL, Clear Channel, MTV and through strategic partner Mediabase, traditional broadcast. Garland is recognised as one of the industry's leading authorities on the intersection of popular entertainment and technology.

#### Acknowledgements

The authors would like to thank Joe Fleischer and Adam Toll at BigChampagne and Gary Eggleton, Chris Carey and Eric MacKay at PRS for Music, all of whom made invaluable contribution to this work. David Touve was especially helpful in providing insights, suggestions and editorial revisions to the work, and for helping bring it towards a conclusion. Finally, a special note of thanks is due to Andrew Bud, Executive Chairman of mBlox, who inspired much of Will's earlier work on the log normal distribution and continues to help the PRS for Music make sense of the unrivalled data sets it has access to.

#### Disclaimer:

This material has been prepared by Will Page at PRS for Music for information purposes only and should not be relied on for any other purpose. It does not constitute the view of the Management or the Boards of MCPS, PRS or any associated company. It is provided for the information of the intended recipient only and should not be reproduced or disclosed to any other person without the consent of the PRS for Music PR department. For further enquiries, information, and to request permissions, please contact: [press@prsformusic.com](mailto:press@prsformusic.com)

## The long tail of P2P

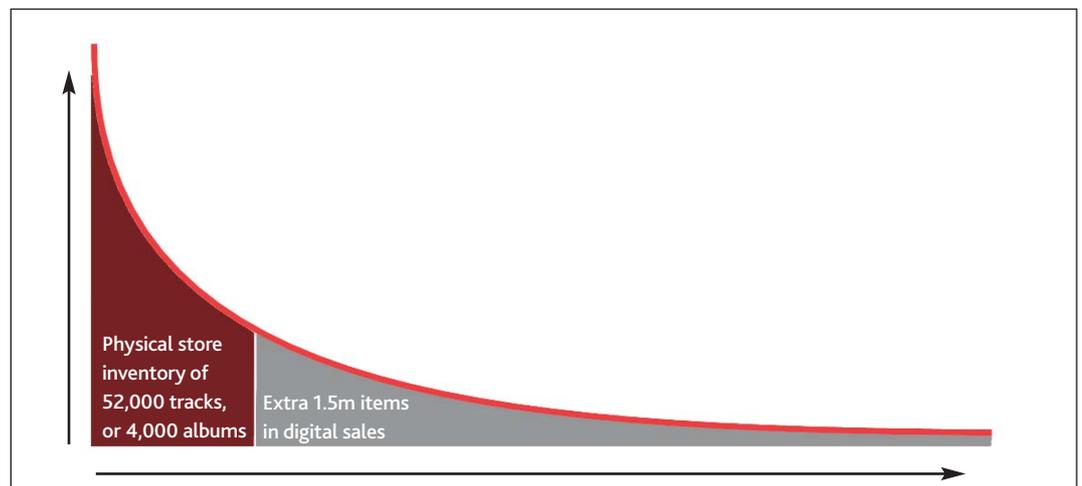
The theory of the Long Tail first came to light in *Wired Magazine* in October 2004, as legal digital music services like iTunes and eMusic were taking off. However, illegal music services like Napster, Grokster or Kazaa had been around, providing digital music fans with a massive choice of music catalogue long before such choice was legally provided. As a result, the well-known anomaly of the digital music world was reinforced - legal services constantly play catch-up with illegal services, and the enforcement of copyright persistently lags advances in technology. With these issues in mind, and the infamous Pirate Bay still very much in the news (and still on folks' desktops), what does the Long Tail distribution profile of hits and niches look like in the world of massive choice that is P2P? Will Page teamed up with BigChampagne's Eric Garland to figure it all out.

The following Insight paper is a bit longer than those previously released by PRS for Music. As such, we thought it best to give the reader, upfront, a sense of where we are headed. In the first section of this Insight paper, we will be releasing the results of a critical inquiry into the music usage patterns within file-sharing networks. Particularly, would a so-called long tail or a pinhead pattern describe the relative popularity of music files within these networks? In the second section of this paper, we will dig into the 'Wherefores' - particularly issues of supply and demand - underlying the usage pattern we found. In the final section, we will consider long-term trends in P2P activity alongside some new behaviours that seem to be emerging. We will then wrap this discussion up with a few final thoughts on the 'paradox of choice'.

#### The Long Tail, a brief introduction

The original Long Tail concept, as laid out by Chris Anderson in a famous *Wired Magazine* article in October 2004, goes like this: if you offer people more choice, and help them make that choice, they will take that choice. It proposed that in a world of widespread Internet access, it no longer makes sense to cater to the public appetite for the most popular CDs, DVDs and books. Instead, even the interests of the smallest

niche might now be served. The lower cost of reaching customers online enables thousands of otherwise unprofitable niches to be profitable. From an economic perspective, this shift in costs would change the 'distribution profile' of transactions in any market to that of "selling less of more." The tail of available niche products would lengthen (supply-side effect) and then fatten with sales (demand-side effect). And so the "Long Tail" emerged.



The theory of the Long Tail took the marketing world by storm, with examples of the theory at work flagged from anywhere and everywhere. Importantly, PRS for Music has been in the Long Tail business since 1921, offering blanket licensing for the entire world's musical repertoire so that both hits and niches might be treated equitably in their supply. Yet, there are demand-side examples of the Long Tail theory not at work in this established Long Tail market. The 1998 Monopolies and Mergers Report on the PRS, for example, described a hit-heavy distribution of blanket licensee revenues. Consequently, doubts have been raised as to how generalisable the examples Anderson had chosen to demonstrate the Long Tail at work would be. These doubts persist given the examples of Long Tail outcomes have often (i) involved unprofitable companies, (ii) referred to volume as opposed to value, and (iii) not recognised the costs of either distributing digital music or processing its granular payments.

Enter Andrew Bud, the Executive Chairman of mBlox and engineering graduate of Cambridge, whose family firm has spent decades understanding the distribution profile of commerce data. Andrew directed the team at the PRS to an influential book by MIT Professor Robert Brown, dating back to 1956, titled *Statistical Forecasting for Inventory Control*. With Mr Bud's help and Brown's methodology, analysts at PRS for Music were able to estimate the demand curve for digital music - of the millions of tracks on the digital shelf, which are moving quickly and

which are collecting dust. To recall, Anderson's theory relies on a change in the nature of the supply curve given barriers to entry fall and a great many new products can now get to the market. However, it takes two curves to tango in economics, and consideration of the demand curve completes the picture.

What we uncovered from that analysis was a shock to some and no surprise to others: a '*hit-heavy, skinny-tail*,' log-normal distribution for legal online music consumption; a distribution not that dissimilar from what one might expect from a more traditional, bricks & mortar store. The Long Tail of digital music commerce had an incredibly lengthy yet dormant tail - more than seventy-five percent of the total inventory of tracks hadn't found a single buyer. Furthermore, a pin-sized head dominated real purchase behaviour. This dormant tail, pinhead pattern appeared across a number of digital music providers, in the markets for singles, albums, as well as streams - the three markets for legally consuming music online.

Intuitively, we were inclined to press pause on our conclusions and ask about illegal music consumption - the black market for music that not only has been with us for years, but also is much larger than its legal equivalent, in terms of numbers of users and available inventory of music (See Box). This is where BigChampagne, with over a decade of experience analysing the data from P2P, stepped in to help produce the following insights.

The 'band to fan' approach embraced by everyone from MySpace to TopSpin is a product of this shift in thinking, courtesy a new set of available tools.

Recorded music businesses are struggling financially, however, and in an 'Attention Economy' (Davenport & Beck, 2001), the future of these businesses remains uncertain. But as WIRED writer Eliot van Buskirk observed just last month: *'[File sharing networks] taught - and continue to teach - valuable lessons to the content industry. Even as music labels and movie studios try to sue peer-to-peer networks out of existence, these same networks have been preparing music labels and movie studios for the emerging social-media world, in which sales form only a small slice of the revenue pie, and what really matters is who likes what, and who pays attention to them.'*

more detailed description can be found in the Appendix to this paper. A simple strap line that captures the overarching message of this work might be: *"everyone gets at least a swap, but the hits are still scooping the pot"*.

Of interest to our analysis was the skew of the distribution for music consumption: does it skew towards the tail, where business is selling '*less of more*'; or does it skew towards the head, where the business of '*more of less*' prevails? To investigate this difference in skew, we opted for a variation on a Lorenz curve, a classic tool in economics for illustrating the proportion of items that drive sales volume. In order to present the data in a manner akin to that of Long Tail aficionados, we have simply flipped the traditional Lorenz curve, to present the cumulative proportion of available tracks across the horizontal axis and the cumulative volume of tracks sold or swapped across the vertical axis. The curves below illustrate (i) what a Pareto-like, 'Long Tail' distribution (red) would be expected to look like, as well as what distributions we observed in (ii) the legal (grey) and (iii) the illegal (burgundy) markets for music online.

### What the Earliest Pirate Markets Taught Us

The Internet market for music began as a pure pirate venue: acquiring the top songs of the day online in 1999 predominantly involved violating copyright laws. Liquid Audio, Rioport, and the eventual arrival of iTunes made the online markets more black and white: pay for the music, or break the law. Today, online music venues tend toward shades of grey. MySpace is legal, but Project Playlist is not quite so legal. Google is legal but SeeqPod is not. QTrax is a 'licensed P2P,' but YouTube features unauthorized music videos. From a consumer perspective, the market is confused and confusing. The earliest pirate markets (Usenet, IRC, and later Napster) provided the first direct observations of the behaviour of music consumers online. The music industry learned to watch and learn. Today, real-time data and business intelligence are music industry obsessions.

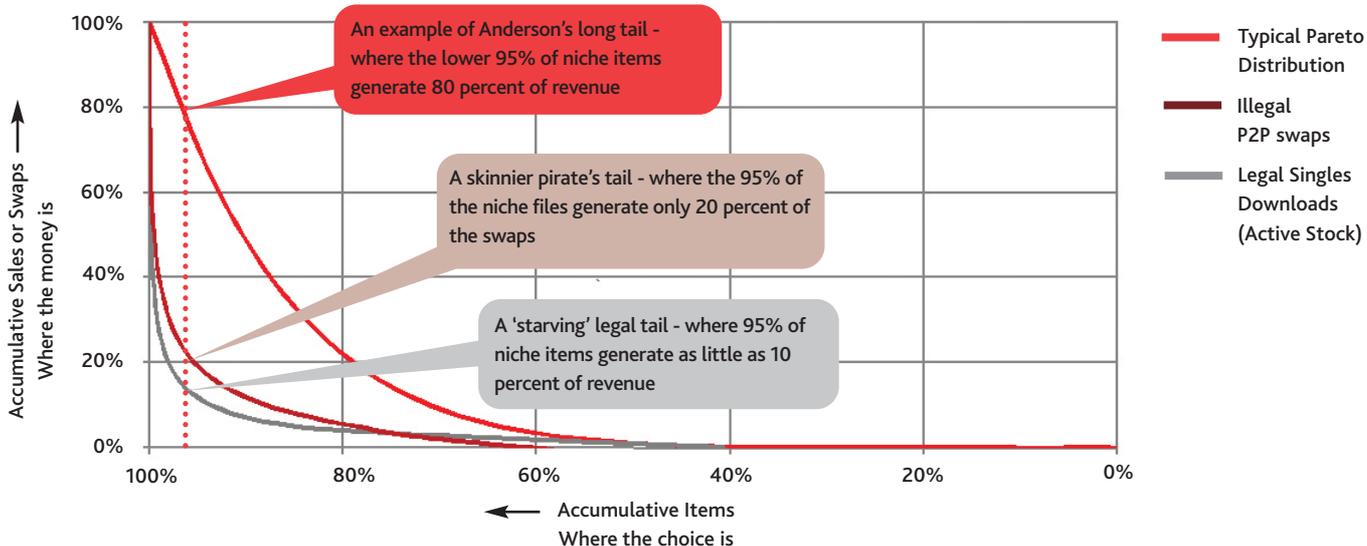
### Deriving a demand curve of P2P

For this research, we collaborated with the folks at BigChampagne to derive some critical understanding of the files swapped on P2P networks. One of the biggest challenges we faced in trying to "pin the tail on P2P" was our need to quantify the universe of music available on file sharing networks. Remember, we were moving from a filtered and controlled legal market, where licenses limit the available tracks, to an uncontrolled market, where few if any constraints are placed upon what music is made available! Put more bluntly, legal music markets offer only what publishers willingly put up for sale, whereas illegal markets swap any file offered by anyone.

At the outset of this research, we did not expect the 'dormant tracks' observation of the legal market to dominate P2P. Instead the data were expected to throw up a tail-biased profile - *a la* the Long Tail theory - due to the unlimited and uninhibited supply of music files. What we found was not what we expected. We will describe the analysis briefly below, while a

## Lorenz curves for legal and illegal music consumption

Source: BigChampagne



### Getting your head around the curves

First, put yourself in the mindset of bottom up, not top down. You are ranking the proportion of total consumption from the least to the most - from right to left. Second, bear in mind that every point on any curve tells you something about consumption in the market described. Each market gets a curve.

For example: On the Long tail curve, your highlighted point tells you that the lower 95% of the inventory would be expected to account

for 80% of consumption (swaps or sales). Conversely, the top 5% of inventory would account for 20% of the consumption.

Sliding down that same curve, and therefore within that same hypothetical Long Tail market, the lower 80% of inventory accounts for 20% of sales - the opposite outlook of the old 80/20 rule. In the P2P market observed by BigChampagne, 95% of the inventory accounted for only 20% of the swaps. Conversely, 5% of inventory accounted for 80% of the swaps. That's not a long tail expectation.

Interestingly and importantly, both legal and illegal music consumption patterns are tucked up against the bottom left axis, contrary to Long Tail-like expectations, suggesting that much of the volume (sales or swaps) is concentrated amongst a small proportion of the available tracks. (Note, if dormant stock is plotted too, the grey line would be more tucked-up.) For economists, the distance between the curve and the bottom-left axis illustrates how fat or skinny the tail of the distribution is. A curve that's tucked up next to the bottom-left border of the chart suggests a skinny tail, or a greater inequality in the distribution of downloads. The closer the curve is to the southwest corner the greater the proportion of volume that is derived from a smaller proportion of the top-ranking items. As the curve moves away from the corner, a lesser proportion of demand is driven by the top-ranking items while a greater proportion is being met by the lower-ranking items. The similarity across both the legal and illegal curves might suggest there's a link, and we'll come to that link in the next section.

### Pondering the Wherefores underlying P2P usage

In the following section, we will offer some of the plausible explanations for this hit-heavy, pin-headed distribution of music-download behaviour we observed on P2P networks. Perhaps it's no surprise that we, being economists, will centre this part of the discussion upon the characteristics of supply, demand and search costs in pirate networks.

### Uninhibited supply

Perhaps the most obvious difference between the pirate and legal markets lay in the nature of supply. In a legal venue (e.g. iTunes, eMusic

or Spotify), the supply of music is often the limiting factor.

Let's take the tried-and-true case of The Beatles, the world's most enduring pop group. The Beatles have yet to appear on the iTunes sales chart quite simply because their music is not available at iTunes at all.

Now consider a pirate venue (P2P file-sharing networks): The Beatles, perennially popular, have been bouncing around the top of the P2P charts for years. And there is no shortage of supply. In fact, every last scrap the Fab Four ever committed to tape is readily available in pirate venues. On some sites, one can download with a single click literally gigabytes of The Beatles' unreleased songs, live recordings, alternate mixes and in-studio chatter.

Within pirate venues for music, supply seems to be virtually limitless. If there were a meaningful audience, however small, for a piece of recorded music, that music is likely to be in supply in the grey market. On P2P networks, however, demand is a very real limiting factor.

### The cost of music search

There is a hidden cost to finding music, what economists would call search costs, and these costs influence demand for music online. In short, people will experiment less when there is a high marginal cost to experimentation. Some of this cost would be in the time we spend searching for music, time that might be better spent elsewhere (like reading this research). Another portion of this cost lay in the risk of buying a dissatisfying song thanks to a poorly informed decision (like a

recommendation from a mate with bad taste in music). If a venue creates (any) marginal search cost (whatsoever), raw demand is inhibited and the tail starts to lose weight.

The very definition of demand is changed when the cost of sampling is nil. In a venue where sampling unknown or little known music has no marginal cost to the customer the Long (Skinny) Tail is more likely to be in full effect. In fact, in the case of some music subscriptions (e.g. eMusic), there is a cost to *not* experimenting. Essentially, a less adventurous subscriber pays a higher, effective per unit price if she does not download the maximum allowable number of tracks in any month.

Searching through millions of music tracks on P2P networks takes time, however - time that not everyone has. As such, behaviour on these networks can still be a function of both online and offline forms of music marketing and recommendation - radio, television, live performances, word of mouth, viral videos, etc. Therefore, after taking into account some geographic differences, the top of the many music charts, from licensed and unlicensed venues, are markedly similar (commercial radio, music television, P2P, iTunes, YouTube, MySpace, even Twitter).

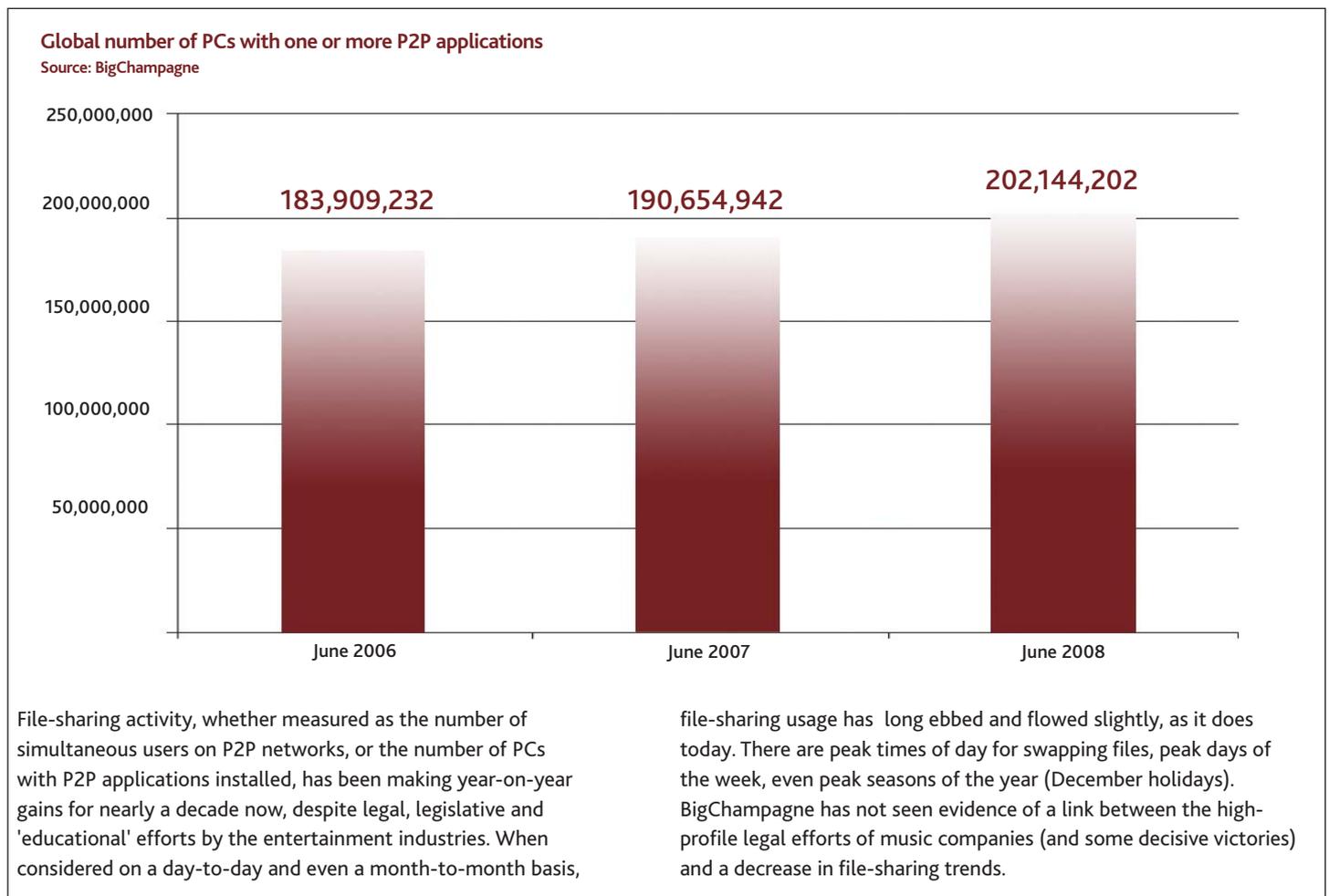
#### Popular music is popular

Since the dawn of the original Napster, the music industry has been

particularly concerned with the Top of the Pirate Pops. More often than not, interest in the pirate market has been focused on the difference between 'what sells' and 'what's stolen.' There are a number of very good reasons for keeping an eye on pirate markets, given these venues mirror and even lead the traditional music market, providing a glimpse of the near-future.

There are many broad and pronounced differences between file-sharing networks and the local record shop, or the iTunes store. But the most popular artists, albums and songs tend to dominate the charts everywhere. In other words: popular is popular. BigChampagne's analysis of the Billboard Top 100 Albums for the last week of April, 2009 reveals that virtually all were available on The Pirate Bay, and most were very popular. These free, unauthorized albums averaged roughly 58,000 weekly downloads each. Lady GaGa's 'The Fame' alone was downloaded 388,000 times in this seven-day period<sup>1</sup>.

Undeniably, the pirate market is a quantifiable expression of the same demand we see at iTunes, or anywhere else. BitTorrent or Limewire popularity of individual titles correlates very highly to sales and therefore (somewhat counter-intuitively) topping the file-sharing charts is generally a harbinger of success in legitimate markets, and has been since the 1990s. In nearly ten years of studying online music, BigChampagne has yet to see a 'big hit,' or wildly popular release in the pirate market that was not also a top seller in the licensed market.



<sup>1</sup>Note that this is just one pirate venue, among many, and that here we are tallying full album downloads only, and excluding individual song downloads.

### The Implications of Pirates and our Principles

In the next section, we will first consider the effect of lawsuits aimed at stemming P2P activity in the context of long-term trends in P2P usage. Second, we will look beyond the Pirate Bay trials and highlight two types of behaviour emerging in both pirate and legal market for music. Finally, we will wrap up this discussion with a few final thoughts.

### Sharpening the stick?

Of course music companies are disappointed to see their greatest sales prospects climbing the pirate charts. There appears to be a devastating effect of file sharing on the music market. What we observe is a music marketplace that once supported first-week sales of as many as two million copies of a number one album (in USA alone) and now serves up albums that top out at considerably less than half that figure.

Ironically, the stubborn popularity of pirate venues has created a cottage industry in anti-piracy efforts - what we might call the business of sharpening the stick in the 'carrot and stick' approach to piracy. Many of these efforts are ongoing, and controversial. However, do we see clear evidence that changes (increases or decreases) in global file-sharing activity can be linked to periods of litigation or legislation? The answer would appear to be, "No."

### After Pirate Bay, what's next?

The last two years brought another rapid transformation in the online music market - a revolution as profound and immediate in some sense as the original onset of MP3 file swapping. What we are seeing is a growth in the nature of music consumption along two seemingly competitive dimensions. These trends lead to a peculiar irony: widespread listening to music that is never stored coincident with vast stores of music to which no downloader ever listens.

The first dimension would involve the *primacy of listening*. In both legal and pirate markets, free streaming of songs, albums and playlists has risen sharply. For example, the most popular songs on MySpace or YouTube now receive hundreds of millions of plays. Similarly, users in file-sharing settings are streaming tracks, particularly within more localised networks. The proportional shift toward streaming presages an ultimate move into 'the cloud' and a de-emphasis on music collection building

Simultaneously, we've noted a spike in a dimension we call *music hoarding*. BitTorrent and one-click hosting sites such as Rapidshare and Megaupload have become increasingly popular as means to acquire music. Both technologies emerged as a function of ever-increasing bandwidth and ever-decreasing storage costs, and both are optimized for large file distribution such as one-click album even 'discography' downloading. Using a Google search for a pirate link, it can be faster and easier to download a full album than to download a single song using Limewire or iTunes. However, a good proportion of music downloaded through these means is *never* listened to, and the files may be ultimately deleted or lost. As such, a sort of hoarding of music files is taking place.

As an activity, only free streaming seems to truly rival MP3 swapping - and taken together, P2P and free streaming now meaningfully compete with traditional broadcast (radio airplay, television) in terms of

impressions. Free streaming might be viewed as being in competition (or tension) with music downloading, or it can be considered the perfect complement. Regardless, the tandem rise of music hoarding and listening primacy signal an ugly potential economic consequence for the recorded music business: the further erosion of legitimate music sales.

### A paradox of choice

Stepping back from this unprecedented study into the shape of demand for music within P2P networks, it's well worth returning to the original driver of this Long Tail debate: the consequences of increased choice. Rigorous analyses of large digital music data sets alongside anecdotal observations are suggesting, contrary to expectations, that in markets offering more choice, the gap between hits and niches appears to be widening. To recall, *PRS for Music* have found hit-heavy, skinny tail distributions in legal digital albums, singles and streams whilst BigChampagne has uncovered a similar distribution for illegal P2P.

Furthermore, there exists evidence that this hit-heavy pattern is strengthening in the booming live music industry. Music promoters comment openly that bands that are four or five down on the bill are relatively worse off than they would have been ten years prior, in real terms. Similarly, pinhead patterns are emerging in TV and film demand, which goes against the original Long Tail logic. According to BigChampagne, torrent downloads of television and film content display an even more concentrated demand for a select few hits than that displayed by music demand. So, while we do see a tail that is long, it is also extremely skinny. The bulk of the business is not in this tail, but instead up near the head, perhaps focused upon an increasingly small bundle of hits. So, if the tail ain't fattening like it should've given so much choice, what's going on?

Resuscitating established research on choice, we find some answers. Anita Elberse from Harvard University pointed to McPhee's "Theory of Exposure," to understand the hit-heavy demand for films from Netflix. According to McPhee's theory, positive feedback effects reinforce the popularity of hits, while dooming items in the tail to perpetual obscurity. In a recent TED lecture, psychologist Barry Schwartz summarizes the arguments within his excellent book, *The Paradox of Choice*: "There is no question that some choice is better than none, but it does not follow that more choice is better than some. There's some magical amount, I don't know what it is but I'm pretty confident that we're long since passed the point where options improve our welfare".

Perhaps however, we're using the wrong language to describe the markets we're trying to understand. References to hits and niches, or heads and tails, might be taking the analysis into an unnecessary corner. There's a temptation to take the anatomy metaphors further, and get hung up on potbellies as well. Andrew Bud, who was responsible for pioneering much of this work, offers another interpretation: focus instead on the gap between rich and poor. Importantly, the relationship between rich and poor can be dynamic. For example, the rich might increase their wealth without the poor getting any less poor, relatively speaking. Ironically, the response to piracy we are avoiding might make some of the richest artists in the world, even richer.

Alternatively, if we economists relax our assumption that humans are rational, self-interested, wealth-maximizing agents we might see ourselves more socially dynamic-humans are also oriented towards social experience. From this sociological perspective there has long existed this concept called "culture" and in its most basic conception, culture would be those things that we share-music, fashion, language and even mindsets. If culture matters, then the widespread popularity of particular music seems less of a paradox. We are innately driven to have some things in common with each other. Furthermore, as David Touve of Vanderbilt University recently suggested, *"P2P networks are such powerful venues for music demand for reasons beyond the economic argument that we can get stuff for free - these networks offer an ideal landscape for sharing the stuff of which culture is comprised."*

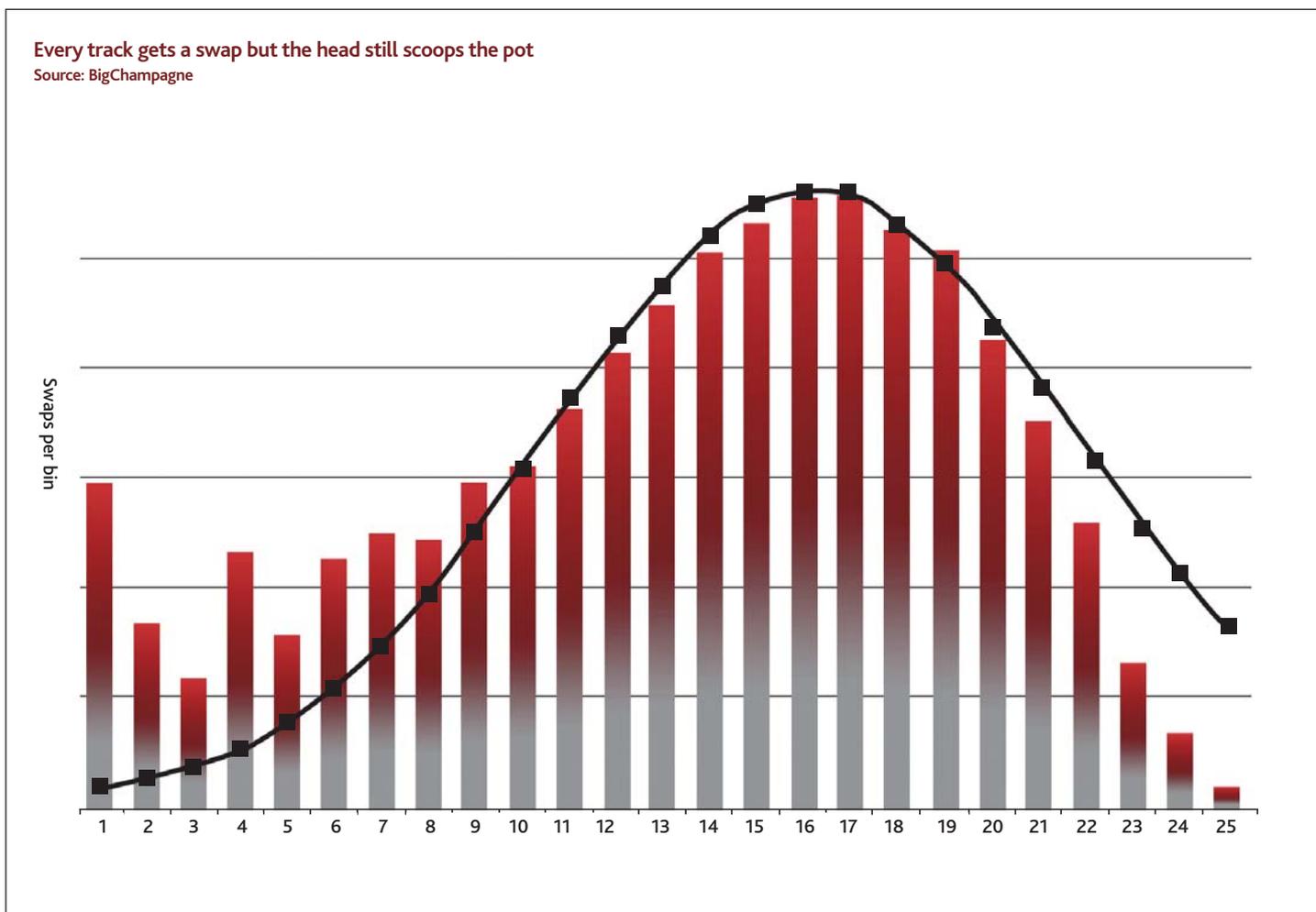
Understanding the markets for music remains the core objective of our research, and P2P networks are now, for better or for worse, part of the music market. But understanding these networks solely as places where opportunities are foregone until these networks are shut down may be clouding our ability not only to rationally assess what is really happening, but also to effectively construct a response. In essence, we may face a certain paradox of choice of our own as we consider our response to these networks. Such a reassessment of the marketplace and our response is an important direction for subsequent research.

## Appendix

### Appendix: Summarising the Statistical Methodology

The biggest task facing the two teams in this analysis of P2P usage data was devising an appropriate structure of log-normal bins, or sales groups, so we could follow Brown's methodology and populate the data set. From the outset, the number of swaps we were dealing with was far greater and more randomly distributed than we had seen in data sets reported from legal music services. Unlike previous analyses of the supply of legal music, the P2P profile gave us no clear indication of the total 'size of the market'. Lacking awareness of those files that were never shared, it would be expected that the number of swaps would be spread more evenly across available tracks, resulting in a more tail-centric profile. Once a suitable analytical structure had been put in place, it was then possible to assign the number of swaps reported to each bin and plot the distribution.

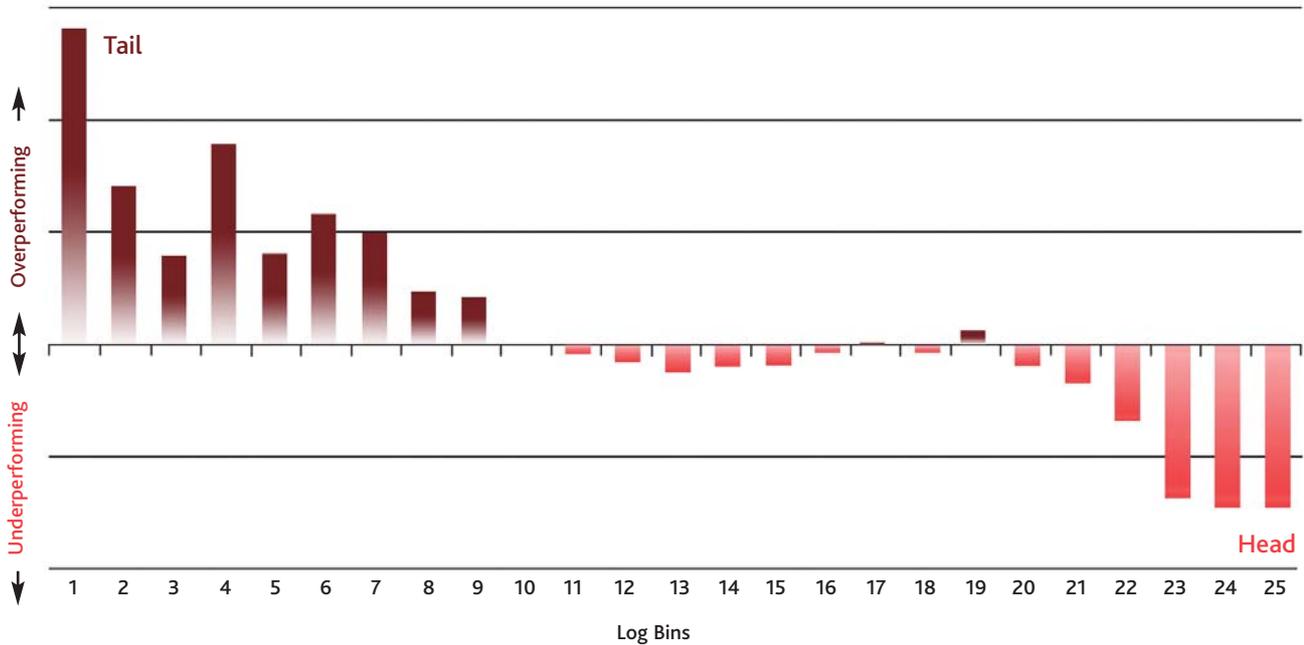
The first chart is a histogram, with the logarithmic bin sizes anonymised. The histogram bars in red represent the number of occurrences of swaps per logarithmic bin structure. What the reader needs to understand is that the groups grow in log-normal intervals, showing how many distinct tracks were swapped 1-to-2 times, then 3-to-6, and 6-to-10 and so on. To clarify, this distribution stretches (left to right) from a bar in the tail with 7 million tracks swapped 1-to-2 times, to the top of the head, a bar with only seven tracks swapped, on average, 13 million times each. The black line is derived from Brown's 1956 methodology, highlighting the expected log-normal distribution given the data we have.



Going back to Brown's 1956 book, history has shown us that when the sales items are plotted logarithmically, a 'Normal' bell-curve distribution often ensues. For any given data set, we would expect some items to sell very little, some to sell an awful lot and the majority to converge around a mean/median.

The second chart is derived from the first, and illustrates, more clearly, how the market performs against Brown's 1956 model. Again, this presentation is akin to asking if the head (or tail) punched above (or below) its expected weight. Recall, Brown's black line is telling us that this is how a market should behave, or the profile of demand should be distributed in a log-normal world. What this chart suggests, however, is the tail outperformed the model, while the head underperformed.

## Evidence of the tail outperforming the head, and the model



As has been seen in every digital music data set that's been plotted to date, the fit of the red line in the first chart is eerily accurate, but with a twist highlighted in the second graph - a clubbed tail and a kick to the head. By that we mean the shape of the demand curve for illegal P2P swaps appears to be that of a log-normal hit-heavy, skinny-tail distribution, with an additional power-law effect in the form of a greater number of tracks are being swapped at least once. In contrast to the legal market, one might view the otherwise dormant tails as being resuscitated on P2P. Put more

bluntly, the 10 million tracks that failed to find a buyer on the legal digital shelf have found (at least) a swapper in this illegal market. Many conclusions could be drawn from these observations, but here's our preferred choice: *If the sellers sell it, it might never be bought; but if the swappers offer it, at least one person will likely take it.* It goes without saying however, that for the creator and artists watching their niche offerings being swapped a single time on P2P, this form of 'freemium' activity may not be paying for lunch.