

Exposure Doesn't Pay the Bills

Artistic Production on Streaming Platforms Under Algorithmically-Induced Audience Uncertainty

Aaron Graybill

10 January 2022

1 Introduction

YouTube, the video streaming giant, reported 19.7 Billion dollars in ad revenue during 2020.¹ Spotify, the music streaming platform, posted a similarly giant 7.8 Billion dollars in revenue from 2020.² Comparing revenues to nominal GDP, this would put YouTube at a similar size to the nation of Afghanistan.³ Revenue numbers like these cannot be ignored.

YouTube and Spotify, among their industry peers like Instagram, TikTok, lie at the intersection of two burgeoning markets: streaming platforms and social media. Services like YouTube, Spotify, Instagram, and TikTok, which I will refer to as *platforms*, provide a rich opportunity for economic analysis.

Platforms are fundamentally a marketplace. *Content creators*, firms, are given an opportunity to connect with consumers who wish to consume their product. However, platforms provide a marketplace unlike anything that has been previously seen. One key distinction between a physical marketplace and a digital platform is that consumption is usually nonrivalrous. One consumer's decision to listen to an artist's newest release on Spotify has no negative impact on another consumer's ability to consume that product. A content creator's inventory is limited only by the bandwidth of the platform and the number of consumers who wish to consume their product. The number of willing consumers is a content creator's audience, and is central to the analysis conducted below.

Audience evolution is another key difference between a streaming economy and traditional market structures. Of course, consumers can seek out content directly if they know that a creator exists, but that does not fully describe consumer-creator matching on a platform. A central component of streaming markets is the *algorithm*. A platform's algorithm can take multiple forms, but I will use it mean a platform's way of analyzing a release by a content creator and deciding which consumers will be shown that release. In some markets, like an outdoor farmer's market, the algorithm is nonexistent. Consumers are required to

¹Available on page 27 at https://abc.xyz/investor/static/pdf/2020_alphabet_annual_report.pdf?cache=8e972d2

²Available on page 5 at https://s22.q4cdn.com/540910603/files/doc_financials/2020/ar/4e770a8c-ee99-49a8-9f9e-dcc191807b56.pdf

³<[https://en.wikipedia.org/wiki/List_of_countries_by_GDP_\(nominal\)](https://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal))> I need a source that isn't wikipedia, but maybe this is stupid anyway.

self-select products they wish to purchase. However, platforms use their algorithm to show content to consumers content they may not have ex ante knowledge.

Content creators have both celebrated algorithms as a way of gaining exposure that would have been impossible in a pre-internet economy. Others have maligned the algorithm for its unpredictability and bias towards content that appeals to the lowest common denominator. However, all successful content creators must acknowledge the importance of a platform's algorithm and adapt to its changes.

A useful anecdote comes from Stevie Wynne Levine, Chief Creative Officer of Mythical. Mythical is a YouTube video conglomerate with over 75 million subscriber across its channels and 25 billion total views on its videos. In reference to how their team analyzes the algorithm, Wynne Levine reveals:

It's something that we do not only every day, but three, four times a day. We analyze each video that we've put out for that day and how we can improve impressions and views. So definitely not something that we set it and forget it, because it's constantly evolving. (para. 12)

The above discussion reveals that content creators have a unique economic problem that is worthy of study. Recent literature has focused on the optimal behavior of consumer and platforms, but content creators have been mostly omitted in the literature of streaming platforms. Work that has tackled artist's optimal behavior has primarily considered fine art in non-digital markets. The analysis below explores whether or not streaming platforms create conditions that mirror fine art markets by concentrating revenue in a small minority of the total number of producers.

In this paper, I will explore how content creators must balance the inherent quantity-quality trade-off present in making content for platform. A higher number of releases offers a content creator more opportunities for the random component of the algorithm to make their content viral. However, investing in greater quality encourages consumers who do know about an artist's content to stream that content more.

I construct a model of artist/content creator behavior in which artists must decide a quantity and quality of art to release in every period. Their quantity choice affords them a number of chances at algorithmic exposure. Additionally, the algorithm will also factor in how many times past audience members consumed an artist's work when deciding audience size. An artist's underlying talent (ability to produce the same quality/quantity at a lower cost) will govern optimal production, and I will show how the distribution of talent relates to the distribution of revenue in successful artists.

I will show that artists early in their career, where new audience members make up a large share of their total audience, are more likely to produce low quality content, and then switch to a low-volume, high-quality strategy later in their career. The interaction of artist decisions in the streaming economy and algorithmic uncertainty is novel to this paper.

The remainder of this paper is broken into (SOME NUMBER) of sections. In the **literature review** I discuss how related literature has informed the expectations of the model. In **the model** section, I hypothesize

results based on the literature, and present the primary model used in this analysis. In the (REST OF SECTIONS GO HERE)...

2 Literature Review

In this paper I will examine how an artist's production decisions and reputation are influenced by algorithmically uncertain audience size in the digital streaming economy. This topic is tethered to multiple lines of research each of which informs the construction of the model below. I discuss connections to the branding literature, the superstar literature, and the novel streaming literature.

The branding literature provides tools to analyze how the information that a firm communicated, its brand, affects consumers' decisions. In this analysis, firms must build a reputation even if that doesn't take the form of a traditional brand, logos, typography, and other factors. The superstar literature analyzes how market concentration can develop, particularly in art markets. I will analyze how the introduction of an algorithm influences the distribution of talent in an art market using the superstar literature as a pre-digitization baseline. Finally, the streaming literature models the incentives and optimal behavior of streaming platforms and end users. The model below supplements the streaming literature by analyzing the artist's problem taking the properties of streaming platform and the end user as given.

2.1 Branding Literature

An artist concerned with growing their audience faces many similar incentives to a new firm trying to develop a brand. One can view a brand as a set of signals about the quality of a firm's product. Branding becomes an important consideration when consumers do not have *ex ante* knowledge about the quality of the product they are purchasing. As such, consumers rely on the signals presented by a firm's brand to inform their consumption decisions. The branding literature begins with the signalling literature pioneered in [Spence \(1973\)](#). In this paper, Spence finds that observable characteristics, both impactful and superficial, can have substantial effects on the hiring decisions of a potential employer. More directly, [Klein & Leffler \(1981\)](#) presents a simple model that identifies key characteristics that a market must possess in order for firms to invest in branding and selling a high quality product. Central to their model is consumer reputation formation. They propose a rather draconian baseline in which consumers' trust can never be regained upon a firm choosing to deceive them. A key finding is that with sufficient differentiation between high and low quality products, some firms may choose to invest in their reputation into perpetuity. This provides evidence that even when a firm's brand offers no intrinsic consumer utility, consumers benefit enough from the information of a brand, that it is worthwhile for both the firm and the consumer to invest in the more expensive branding. [Shapiro \(1983\)](#) generalizes the model proposed by Klein and Leffler to a case where reputation can exist on a continuum of values and can evolve in a less austere manner. Shapiro confirms the results of Klein and Leffler while expanding on the fragility of branding. Shapiro finds

that even when a firm is able to charge more for a branded product than an unbranded alternative, this premium is fleeting and sensitive to changes in consumer preferences.

The application of the branding literature to the problem posed in this paper lies in the way that streaming platforms' algorithms reveal information and content to consumers. Generally streaming platforms allow users to choose their own content, but many platforms like YouTube, Spotify, and Apple Music also algorithmically provide content based on the consumer's history of the content they have consumed and how they have consumed it (number of times, shares, etc.). When this content is previously unknown to the consumer, we can view the platform as relying on its branding to present desirable content to the user. I will extend the branding literature by exploring the context in which reputation is subject to uncontrollable shocks that can positively or negatively influence next period's reputation.

2.2 Superstar Literature

A related strand of literature to branding is the superstar literature which lies more towards cultural economics. Rosen, the preeminent author the superstar literature, characterizes a superstar market as having a: "relatively small numbers of people [who] earn enormous amounts of money and dominate the activities in which they engage," (Rosen, 1981, p. 845). (LET ME KNOW IF THAT CITATION IS DONE WRONG) In his paper, he presents a model where consumers can perfectly observe talent ex ante. In this framework, Rosen finds that small increases in underlying artistic talent can have disproportionately large increases in resultant revenue in market equilibrium. This leads to revenues being concentrated among only a few artists. Art market concentration is further explored in MacDonald (1988) which takes an alternate modeling approach and introduces uncertainty in the talent of an artist. This uncertainty is market-wide where neither the consumer nor the artist knows their talent until they have performed. This model also intersects with the branding literature because it explores how an artist's perceived talent evolves over multiple periods based on the quality of their performance.

The superstar literature is important to this analysis because it emphasizes the artist's production decisions and how they affect market revenues. As I will discuss later, much of the recent literature on streaming economies has focused on the streaming platform and the end consumers, so the superstar literature gives a more nuanced view of how artistic production can be modeled and optimized. However, I add to the superstar literature by modernizing the analysis and seeing if the same patterns emerge under a digitally-based economy. In particular, I will provide further insight into how the distribution of underlying talent may or may not be reflected in the realized distribution of talent in successful artists.

2.3 Streaming Literature

The third and most closely related field of study is the streaming literature. The advent of streaming has garnered considerable attention from both theoretical and empirical angles. Streaming platforms can take

many forms, but I will follow the characterization given in [Thomes \(2013\)](#). He characterizes a streaming economy as an internet-based two-sided media marketplace. The streaming platform is in the middle of this two sided marketplace. The first side of the marketplace is the streaming platform accepting media from content creators. The other side of this marketplace is the streaming service delivering this content to consumers. Usually money is changed hands on both sides of this market. Researchers have focused on various aspects of the streaming economy as it has evolved. I will begin by giving an overview of some of the key topics addressed by theoretical papers. After discussing theory, I will mention some relevant empirical studies that analyze how digitization has affected the distribution of artist popularity—a central question of the research at hand.

Early papers in the streaming literature investigate how streaming may curb pirating, the (usually free) illegal download of unlicensed music, like on Napster. One such early paper is [Thomes \(2013\)](#) which takes the perspective of the music streaming platform when deciding how to price its paid subscription service relative to its free-with-ads alternative. This paper takes the artist's behavior as given, and does not tackle the conditions under which artists will choose to produce content for the platform.

Another more recent paper that takes the streaming platform's perspective is [Bender et al. \(2021\)](#). This paper analyzes competition between permanent digital MP3 sales and streaming platforms. It analyzes consumer demand, and how the platform should optimally set its royalty to attract artists to the platform. The authors find that it is the most popular artists that might choose to hold out and sell their work only via permanent download, a result seen anecdotally with the Beatles who kept their music off of streaming services for a famously long time.⁴ The analysis below examines how artists should optimally produce once they have already committed to making content for a streaming platform.

[Hiller & Walter \(2017\)](#) incorporates elements from both [Thomes \(2013\)](#) and [Bender et al. \(2021\)](#) by modelling an economy with digital purchases in addition to free and paid subscriptions on a streaming service. The authors investigate the artist's decision to produce one high-quality piece of art versus multiple comparatively lower-quality pieces. The authors find that the streaming economy fosters an environment in which profit-maximizing artists focus on generating high-quality singles instead of lower-quality albums. I will examine this quantity-quality trade-off will below, but with the key difference of uncertain audience size. The proceeding analysis acknowledges the fact that releasing more content may increase an artist's chance of reaching a larger audience.

Another driving question of this research pertains to the debate of whether or not streaming platforms have created a “long tail” of products. I will now summarize some empirical work that analyzes this topic.

The term long tail was popularized in [Anderson \(2006\)](#). The principle of the long tail is that digitization of commerce allows consumer's access to a wider variety of products—far more than a brick and mortar store could every stock. The greater variety of products allows consumers to pinpoint the product that best suits their preferences. A diverse product set results in many products having a small number consumers.

⁴A discussion of the Beatles decision to stream their music is available here: <https://www.fastcompany.com/3054965/its-official-the-beatles-are-coming-to-spotify-apple-music-and-more>

As such, the distribution of consumers per product should have much longer tail than pre-digitization. The streaming economy is an excellent example of where this long tail might exist as Spotify can stock hundreds of times more songs into a server than a vinyl record store could ever stock in-store. The long tail hypothesis would suggest a diffusion of revenues across many artists. However, [Rosen \(1981\)](#), finds that revenue should be concentrated among only a few artists. These factors are not necessarily at odds, we may have most of the revenue while still having a long tail, but this paper will examine whether or not we have a long thick tail, or an initial bump with a long thin tail afterwards.

That being said, opinions are mixed as to whether or not the long tail hypothesis holds empirically. [Elberse \(2008\)](#) initially pushed back against the idea of a long tail. She cites data from Quickflix (a now-defunct Australian pay-per-view movie streaming platform) which showed that a small number of DVDs comprised a large portion of sales. The proceeding analysis in this paper will model a different type of streaming platform, one in which customers have unlimited access to content either for free or for a monthly subscription fee.

A more recent alternate perspective on the long tail, [Aguilar & Waldfogel \(2018\)](#), posits digitization substantially lowers the entry cost of new firms. Lower entry costs allow firms with lower expected profit into the market, increasing the diversity of sellers. The authors robustly show that consumers value having a variety of producers when it is difficult to forecast an artist's talent before purchasing. They argue that the value of variety provides further justification that there exists a long tail of producers. I will examine whether or not producers naturally form a long tail even under uncertainty about the number of consumers that they will be able to reach with their product.

2.4 This Paper's Contribution to the Literature

The preceding discussion reveals a few key unexplored area that this paper will investigate. One, previous literature has focused on behavior of the streaming platform and the end-user. This analysis contributes to the comparatively-understudied role of the artist in the streaming economy. Two, this paper will endogenize audience development, a similarity to the branding and superstar literature not yet applied to the streaming literature. Third, this paper will contribute the debate of the long tail in the streaming literature by exploring the equilibrium distribution of artist talent. Finally, This paper will also contribute to the presently-unexplored role of unpredictable algorithms on streaming platforms in shaping an artist's career and optimal behavior.

3 The Model

The streaming economy has three primary actors: the end consumer, the streaming platform, and the artist. For the purposes of this analysis, I will assume that the behavior of the consumer and the streaming platform are exogenous. I will limit attention to the artist who must choose a quality, z_t and quantity, m_t of

art to produce in every period. The streaming platform’s algorithm is driven by consumer engagement last period, n_{t-1} . The algorithm combines past user engagement with the number of new releases to decide the number of new consumers to show an artist’s work to. I will assume that consumers have no ex ante knowledge of an artist and require the algorithm to reveal an artist to them. I will call these algorithmic revelations, impressions and denote them I_t . Once a consumer has received an impression, their number of streams will depend on the quality of the art that the artist released in that period. If a consumer has received an impression, they are under no obligation to stream, the number of streams is fully determined by the quality of the art. Impressions only serve as a way for a consumer to be made aware of an artist’s portfolio. After the consumer decides to stream, the algorithm observes the streams per impression and uses this to decide how many times to show the artist’s work next period.

The artist earns a royalty every time a song is streamed, and the artist tries to maximize this discounted flow of royalties. The trade-off between quantity of releases and the quality of those releases will be central to their maximization problem. Increases to quality ensure that once a consumer receives an artist’s work, they will consume that product more. Increasing quality also has the inter-temporal benefit of encouraging next period’s algorithm to show the art to more consumers. On the other hand, the artist can increase their quantity at the expense of quality. The algorithm will have more pieces to show to consumers (increasing the probability that a given consumer discovers an artist), but it also leaves audience size more up to the random component of the algorithm. I will not model the algorithm in any more detail than the components above. The algorithm, in this analysis, will serve only to link the number of impressions this period to the number of impressions as given. Further analysis is required on a streaming platform’s optimal algorithm construction, but that is outside the interests of this paper.

3.1 Hypotheses

Before describing the model in detail, I will first lay out hypothesized results based on the literature presented above. The first hypothesis pertains to the “long tail” theory described in [Aguilar & Waldfogel \(2018\)](#). Following [Rosen \(1981\)](#), I will measure the long tail of outcomes by examining the convexity of the expected profit function in an artist’s underlying talent. I suspect that the imperfect information present on streaming platforms will make it harder for consumers to find and stay with artists that best fit their preference which may weaken the market power of the most talented artists. I pose the following:

Hypothesis 1: Unpredictability of a streaming platform’s content matching algorithm will decrease the convexity of profits in talent.

The second hypothesis relates to audience development. Similar to the work of [Bender et al. \(2021\)](#), I will examine how the choices of established artists differ from artists with smaller starting audiences. In a one period model, smaller artists do not carry a large audience from the previous period, so a high quality product does not get shown to enough consumers to justify a high quality strategy. As such, I propose:

Hypothesis 2: Artists with smaller initial audiences are more likely to choose a low-quality, high-quantity strategy relative to established artists.

The final hypothesis pertains to how the previous hypothesis is affected by time. When there are multiple periods, a relatively new artist can rely on producing a high quality product knowing that the future benefits of high quality products will outweigh the short lived and uncertain benefits of a low quality product. I propose:

Hypothesis 3: Artists will prioritize quality more in a multi-period setting than they will in a one-period environment.

3.2 Consumer Behavior

I assume there is an infinitely large market of consumers with identical preferences on a streaming platform. This assumption does not fully depict the consumer-base of a streaming platform. However, for an emerging artist in an established genre, this assumption is more realistic. To a new artist, the number of potentially-accessible consumers can be almost limitless. Further, if an artist enters an existing genre, there is a clear sense of what is successful across the entire market, so there is some observable homogeneity in consumer preferences.

Ex ante, consumers have no knowledge of a given artist and require the algorithm to reveal an artist to them. Upon receiving an impression of an artist, the consumer gains knowledge of all of an artist's work from that period, not just the piece they were exposed to. I will assume that after an impression a consumer will stream the artist's work $n(z)$ times, where z is the quality of the art. I will assume that $n_z > 0$ and $n_{zz} < 0$, so increases to quality always increase demand, but at a lessening rate. A notable assumption in the above modelling decision is that the quantity of releases has no influence on the amount of streams by a given consumer. Consumers see the quality and might choose to stream one song all n times, or they might spread their consumption across multiple pieces of art.

3.3 Artist Behavior

The artist's problem is to maximize their discounted stream of profits over multiple periods. The artist generates revenues from per-stream royalties and faces a cost according to their quality and quantity. In particular, I will assume that each stream earns the artist an amount r , so the total revenue in each period is the total number of streams time r . The artist chooses to produce m pieces each of which at the same quality z . I will assume that the artist's combination of m and z also incur a cost $C(m, z; \kappa)$ where κ is the artist's underlying talent that makes production easier. I will impose some standard assumptions on C . Namely, I will assume $C_m > 0$, $C_z > 0$, $C_{mm} > 0$, $C_{zz} > 0$, $C_{z\kappa} < 0$, $C_{m\kappa} < 0$. The assumptions say that increases to quality or quantity also increase cost, and additional units of quantity of quality are more

costly than the previous. The cross partial derivatives say that increasing talent decrease the marginal cost in each of the inputs. In order for κ to be meaningful as a talent parameter, we should require that at every input, an additional unit of talent makes the next unit of production less costly. I will further assume $C_\kappa < 0$ and $C_{\kappa\kappa} > 0$, so increases to underlying talent lower costs, but additional increases to talent are less and less impactful. The only sign without an immediate sign choice is the quantity-quality cross partial derivative, $C_{mz} = C_{zm}$. A positive cross partial derivative says that a one unit increase in quality increases the marginal cost of an additional unit of quantity. We can also interpret this as a one unit increase in quantity increases the marginal cost of an additional unit of quality. More broadly, the sign of this partial derivative determines whether or not the inputs are complementary. For this analysis, I will assume that $C_{mz} \geq 0$. This assumption is reasonable considering that an artist only has finite time in every period, so every minute they spend on using more of one input, decreases the time they have for the other.

3.4 Audience Evolution and Algorithmic Behavior

First, a distinction between audience and streams. For the purposes of this analysis, an artist's audience at time t , A_t , is the number of consumers who are aware of an artists work. Again, these audience members are under no obligation to stream, but they will be able to observe an artists quality to inform their consumption decisions.

I impose three heuristics on how an artist's audience evolves over time. First, some proportion of last period's audience should be retained between periods. Second, a streaming platform should have an algorithm to determine the number of new audience members that an artist has. Finally, there is some unforecastable random noise present in how the algorithm behaves, at least in the eyes of the artist.

I will first assume that every period, a fraction δ of the audience "forgets" about an artist and needs reimpression in order to consume an artist's work again. As such, the share of last period's audience that endures is $(1 - \delta)$. We can think of $1 - \delta$ as the share of audience members who naturally remember an artist or as the fraction of consumers who are forced to remember an artist from reminders by the streaming platform.

Now I will characterize how new consumers can be added to an artist's audience. I assume that the streaming platform's algorithm governs the number of new impressions for each song released by an artist. For the algorithm to be a meaningful tool, it should not be entirely random. It should use some measure of engagement to dictate how many impressions in the next period. For the purposes of this model, the algorithm will use last period's number of streams per audience member as the measure of engagement. By construction, at time t , the algorithm will use $n(z_{t-1})$, so the algorithm is indirectly incorporating quality.

There are multiple ways to interpret the uncertainty in the algorithm. One such way is to interpret a streaming platform's algorithm as an imperfect instrument that measures talent. An alternate way to interpret algorithmic uncertainty is in the context of producer uncertainty. In this case, the artist understands the average effects of the algorithm, but the streaming platform intentionally or inadvertently obfuscates

exactly how the algorithm behaves so there is always some artist uncertainty about the true number of impressions in the next period.

Tom Scott (2017) gives an engaging look into why YouTube's algorithm is unpredictable. He notes that when a streaming platform reveals information about a streaming platform's algorithm, get-rich-quick content creators produce inferior content that is only intended to exploit these trends. He also notes that many algorithms are constructed using machine learning and huge training data sets and have too many inputs to be easily understood.

An additional concern when modeling an uncertain algorithm is that artist with higher talent may be favored by the random component of the algorithm. However, most artists, at least for professional content creators, can observe any information in the algorithm that might benefit them when making their product. By definition, the most algorithmically favored products are shown the most. As such, I will assume that artists have ample opportunity to analyze successful content and arbitrage away any useful information. This assumption says that there are no frictions in obtaining information about the algorithm. Certainly dedicated media teams have a better ability to spot algorithmic trends than single producers, but this relationship is outside the scope of this paper. Further analyses could explore this relationship in more detail, but this model will assume that talent and algorithmic uncertainty are independent.

With the aforementioned assumption, I construct the algorithm as follows. First denote the impression algorithm for an artist's i th art piece in period t as $I(n_{t-1}) + \varepsilon_{it}$ where ε_{it} is a mean zero independent identically distributed random variable. Both I and ε have units of number of additional people in audience. Assuming the algorithm is a useful, if imperfect, measure of quality, I will assume that $I_n > 0$. So increases to engagement increase the expected number of impressions per release.

For each art piece that an artist produces in the present period, they must submit this work through the algorithm which is subject to random noise. If the artist releases m pieces in a given period, the total number of impressions from the audience is then given by $mI(n_{t-1}) + \sum_{i=1}^m \varepsilon_{it}$. As such, the expected number of impressions is simply $mI(n_{t-1})$ with variance $m\text{Var}(\varepsilon)$. Increasing the number of releases increases the expected audience, but equally increases the variance of outcomes. Putting all of these pieces together, the equation of motion for audience size at time t , A_t , is given by:

$$A_t = (1 - \delta)A_{t-1} + mI(n_{t-1}) + \sum_{i=1}^m \varepsilon_{it} \quad (1)$$

We can visualize audience development as follows:

4 The Artist's One-Period Maximization Problem

We can assemble the above pieces into the respective one-period revenue maximization problem. The timing of the model is as follows, the artist chooses m and z , then the random variables in the algorithm

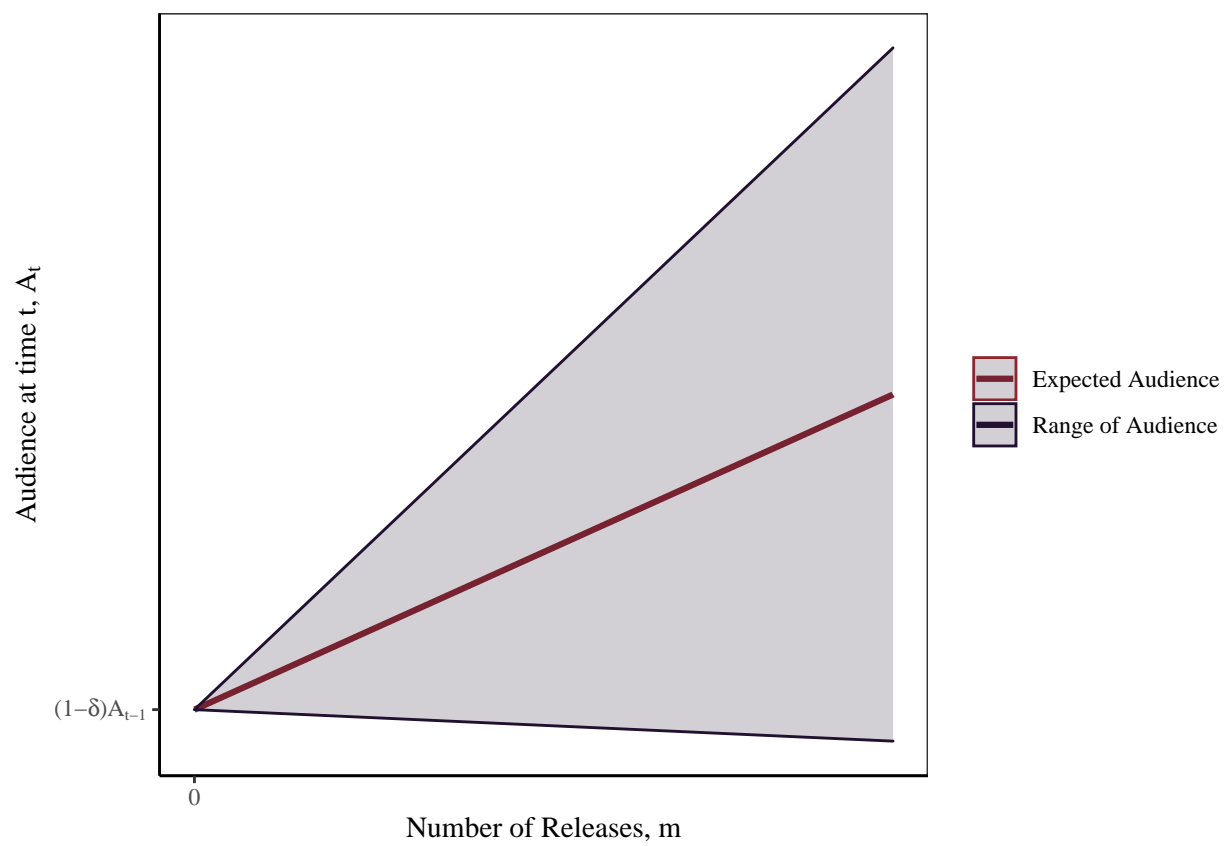


Figure 1: Increases to quantity produced increase expected audience and variance of outcomes

are realized, then consumers stream the artist work according to $n(z)$. In the case where there is only one period, we can interpret n_{t-1} as the total amount of reputation that an artist has earned over the duration of their career up to now. Similarly, A_{t-1} becomes the entire audience that has been retained. In a one period model, objects in the equation of motion represent a lifetime's work, not just one period.

As such, we can express the artist's problem as:

$$V(m, z) = \max_{m, z} \left\{ E [r A_t n(z) - C(m, z; \kappa)] \text{ s.t. } A_t = (1 - \delta) A_{t-1} + m I(n_{t-1}) + \sum_{i=1}^m \varepsilon_{it} \right\} \quad (2)$$

Substituting in with A_t allows us to solve this problem as an unconstrained maximization problem with the following first order conditions:

$$\begin{aligned} \frac{\partial V}{\partial z} &: r [(1 - \delta) A_{t-1} + m I(n_{t-1})] m n'(z) - C_z(m, z; \kappa) = 0 \\ \frac{\partial V}{\partial m} &: r I(n_{t-1}) n(z) - C_m(m, z; \kappa) = 0 \end{aligned} \quad (3)$$

I will now use the implicit function theorem to interpret some comparative statics in terms of the parameters of the model. I summarize the results below

Table 1: Signs of first order conditions for the one-period model

X	$\partial m / \partial X$	$\partial z / \partial X$
r	–	–
δ	0	+
A_{t-1}	0	–
n_{t-1}	–	–
κ	–	–

The first column gives the relationships between increasing the parameters and the optimal choice of quantity, m . Producing more quantity incurs costs that the artist substitutes away from when given more slack by the other parameters. Interestingly, the previous period's audience makes no impact on the artist's choice to produce a greater quantity. This comes from the fact that prior audience is sunk in m . Since m can only influence the number of new listeners, the artist need not consider how much of the previous audience they've retained.

Since z does not affect the future engagement measures (because I only consider one period), the artist will try to substitute away from using z in production because it is more costly. As such, when the royalty rate increases, the artist will be able to recoup the same amount of revenue selling fewer units, so they will chose to lower their z due to it's increasing costliness.

4.1 One Period-Binary Choice Example

I will now explore the case in which the artist can only choose from one of two options. In order to collapse the problem to a binary choice set, instead of using a cost function, I will add an additional budget constraint, $C(m, z) = Y$ which implicitly defines z in terms of m . As such, let's consider the case where the artist is choosing to produce either 1 or 2 products. If the artist chooses $m = 1$, they can produce one higher quality good at quality \bar{z} . In the other case, the artist can produce 2 goods, each at quality \underline{z} . The artist will then choose:

$$\max \{r((1 - \delta)A_{t-1} + I(n_{t-1})) n(\bar{z}), r((1 - \delta)A_{t-1} + 2I(n_{t-1})) n(\underline{z})\} \quad (4)$$

The artist's optimal production choice can be summarized as

$$\begin{cases} \underline{z} & n(\bar{z}) < \left(1 + \frac{I_0}{I_0 + (1 - \delta)A_0}\right) n(\underline{z}) \\ \text{Either} & n(\bar{z}) = \left(1 + \frac{I_0}{I_0 + (1 - \delta)A_0}\right) n(\underline{z}) \\ \bar{z}, & n(\bar{z}) > \left(1 + \frac{I_0}{I_0 + (1 - \delta)A_0}\right) n(\underline{z}) \end{cases} \quad (5)$$

By monotonicity of demand $n(\bar{z})/n(\underline{z}) > 1$, so we can think of $\frac{I_0}{I_0 + (1 - \delta)A_0}$ as the minimal premium for which the artist will produce the high quality option. Begin by noting that the denominator in the above expression is the expected audience size when the artist chooses the high quality option. As such, we can interpret the premium as the percent of expected audience that is earned by the algorithm. Holding other factors equal, an artist with a larger initial audience, A_0 , is more likely to have a smaller premium to induce high quality production relative to a new artist where most of their audience is new. Therefore, established artists are more likely to produce high quality content. In contrast, new artists are likely to produce a greater number of lower-quality art and rely on the algorithm to get their art into the hands of new consumers.

4.2 One Period with Talent

Since the talent parameter κ enters through the cost function, it is not represented in the analysis above. I will now modify the model slightly, allowing κ to enter into the binary choice problem. Instead of constraining the artist to produce only one or two items at qualities \underline{z} or \bar{z} , I will now examine the case in which the artist still chooses to produce an items of qualities \underline{z} or \bar{z} , however the artist's talent dictates the number of releases they can produce. In particular, let $\underline{m}(\kappa)$ be the number of low-quality releases that an artist of talent κ can produce in one period. Similarly, define $\bar{m}(\kappa)$ be the number of high quality releases the same artist could produce in one period. In the eyes of a given artist (who only has one talent), they only have two production options, however artists with differing talents face different (though still binary) production options.

The functions $\underline{m}(\kappa)$, $\overline{m}(\kappa)$ should mirror reality in their properties. Namely should require that for all κ , $\underline{m}(\kappa) > \overline{m}(\kappa)$, so more low quality products can be produced than high quality ones. Furthermore, both $\underline{m}(\kappa)$ and $\overline{m}(\kappa)$ should be increasing in κ and should have diminishing marginal product, $\overline{m}''(\kappa), \underline{m}''(\kappa) < 0$.

We can formulate the artist's one-period maximization problem as:

$$\max \{r((1-\delta)A_{t-1} + \overline{m}(\kappa)I(n_{t-1}))n(\overline{z}), r((1-\delta)A_{t-1} + \underline{m}(\kappa)I(n_{t-1}))n(\underline{z})\} \quad (6)$$

With optimal choice of quality satisfying:

$$\begin{cases} \underline{z} & n(\overline{z}) < \left(1 + \frac{(\underline{m}(\kappa) - \overline{m}(\kappa))I_0}{\overline{m}(\kappa)I_0 + (1-\delta)A_0}\right) n(\underline{z}) \\ \text{Either} & n(\overline{z}) = \left(1 + \frac{(\underline{m}(\kappa) - \overline{m}(\kappa))I_0}{\overline{m}(\kappa)I_0 + (1-\delta)A_0}\right) n(\underline{z}) \\ \overline{z}, & n(\overline{z}) > \left(1 + \frac{(\underline{m}(\kappa) - \overline{m}(\kappa))I_0}{\overline{m}(\kappa)I_0 + (1-\delta)A_0}\right) n(\underline{z}) \end{cases} \quad (7)$$

Proposition 1. *Increases to talent incentivize an artist to produce low quality work whenever their talent satisfies $\underline{m}'(\kappa)\overline{m}(\kappa) - \overline{m}'(\kappa)\underline{m}(\kappa) \geq 0$.*

Proof. By 7 the low quality strategy becomes more desirable when the following quantity increases:

$$\left(1 + \frac{(\underline{m}(\kappa) - \overline{m}(\kappa))I_0}{\overline{m}(\kappa)I_0 + (1-\delta)A_0}\right)$$

Therefore, the low quality strategy gains desirability from increases to κ whenever the derivative of the above expression is greater than zero. The κ derivative is:

$$\frac{(\underline{m}'(\kappa) - \overline{m}'(\kappa))I_0 (\overline{m}(\kappa)I_0 + (1-\delta)A_0) - (\underline{m}(\kappa) - \overline{m}(\kappa))I_0 (\overline{m}'(\kappa)I_0)}{(\overline{m}(\kappa)I_0 + (1-\delta)A_0)^2}$$

The denominator is always positive, so positivity on the derivative only requires the numerator to also be positive. Expanding and simplify that expression gives the following inequality:

$$(\underline{m}' - \overline{m}')(1-\delta)A_0 + (\underline{m}'\overline{m} - \underline{m}\overline{m}')I_0 > 0$$

Now suppose that $(\underline{m}'\overline{m} - \underline{m}\overline{m}') > 0$, and by construction $\underline{m} > \overline{m}$. Note that for an arbitrary $a, b, x, y \in \mathbb{R}^+$, when $ax - by > 0$ $x > \frac{b}{a}y$. The condition $x > y$ is weaker than and is implied $x > \frac{b}{a}y$ whenever $a > b$. Letting $a = \overline{m}, b = \underline{m}, x = \underline{m}', y = \overline{m}'$, shows that both terms in the above inequality must be positive, so the whole derivative must also be positive. \square

This sufficient condition $\underline{m}'(\kappa)\overline{m}(\kappa) - \overline{m}'(\kappa)\underline{m}(\kappa) \geq 0$, can be equivalent to: talent causing a 1% increase in the number of high quality units available for production results in a greater than 1% increase in the number of low quality goods available for production.

This says that increases to talent push an artist towards low quality production whenever low quality production is cheap. We should not innately have a sense for whether or not $\underline{m}'(\kappa)\overline{m}(\kappa) - \overline{m}'(\kappa)\underline{m}(\kappa) \geq 0$ holds in reality. One could argue that an additional percent of low quality production may be harder because you produce more low quality goods, so additional units may be more difficult. Alternately, high quality production may be harder because these goods are inherently harder to produce.

5 Two Period Maximization

I now explore the case where the artist has the same options as in the one-period no-cost case, presented in 4.1. The artist can choose one or two products at quality \bar{z} and \underline{z} respectively. However, in this case, I examine the artist's behavior over two periods. In this case, their choice of z in the first period influences the number over impressions that the algorithm produces next period. I will now introduce some new notation. Let \bar{I} be the impressions awarded to the artist when their choice of quality \bar{z} induces consumption $n(\bar{z})$. Define \underline{I} analogously. I define discount factor β that deflates the value of profits in subsequent periods.

The artist has four options, either quality in either period. I will denote the items in the strategy set: $\{\underline{z} \rightarrow \underline{z}, \bar{z} \rightarrow \underline{z}, \underline{z} \rightarrow \bar{z}, \bar{z} \rightarrow \bar{z}\}$, where $\underline{z} \rightarrow \bar{z}$ is the strategy with low quality in period one, and high quality in period two and so on. (HOW AWFUL IS THAT NOTATION?) Each strategy has the following utility.

$$\begin{aligned}
E[V(\underline{z} \rightarrow \underline{z})] &= r((1-\delta)A_0 + 2I_0) n(\underline{z}) + \beta(r((1-\delta)A_1 + 2I(n(\underline{z}))) n(\underline{z})) \\
E[V(\bar{z} \rightarrow \underline{z})] &= r((1-\delta)A_0 + I_0) n(\bar{z}) + \beta(r((1-\delta)A_1 + 2I(n(\bar{z}))) n(\underline{z})) \\
E[V(\underline{z} \rightarrow \bar{z})] &= r((1-\delta)A_0 + 2I_0) n(\underline{z}) + \beta(r((1-\delta)A_1 + I(n(\underline{z}))) n(\bar{z})) \\
E[V(\bar{z} \rightarrow \bar{z})] &= r((1-\delta)A_0 + I_0) n(\bar{z}) + \beta(r((1-\delta)A_1 + I(n(\bar{z}))) n(\bar{z}))
\end{aligned} \tag{8}$$

5.1 Numerical Simulation of Two Period Optimal Behavior

The optimal strategy is the maximum of the four strategies listed in 8, but the conditions under which certain strategies are preferred are hard to characterize. I will use numerical simulation to describe the conditions that produce certain optimal strategies.

In order to simulate the possible conditions, I generate sample values for each of the variables in 8. Absent a more informed prior, I uniformly sample each variable in a reasonable range of values. I summarize these ranges below:

Table 2: Sample range for simulated parameter values. *Indicates that the maximal value is limited by the high quality alternative

Variable	Symbol	Minimal Possible Value	Maximal Possible Value
Streams per capita, high quality	$n(\bar{z})$	0	10
Streams per capita, low quality	$n(\underline{z})$	0	10*
Starting audience	A_0	0	100
Discount factor	β	0	1
Audience depreciation rate	δ	0	1
Expected impressions per song, high quality	$I(n(\bar{z}))$	0	100
Expected impression per song, low quality	$I(n(\underline{z}))$	0	100*
Initial expected impressions	I_0	0	100*

The values above are primarily aimed at describing the behavior of new artists, where the algorithm will not show their content to more than 100 new consumers per release. Furthermore, this assumes that consumers will stream an artist's portfolio no more than 10 times, an assumption that reflects more durable entertainment like a YouTube video instead of a song. Additionally, I filter out any realization of the random variable in which I_0 or $I(n(\underline{z}))$ are greater than $I(n(\bar{z}))$ because these are cases in which the algorithm rewards new and low quality art more than high quality art. If the algorithm is to be meaningful, then it should prioritize high quality content. Similarly, I filter out any cases when $n(\underline{z}) \geq n(\bar{z})$ because high quality work should be streamed more than low quality alternatives.

With that, I generate a simulation dataset consisting of a realization of each of the eight variables in the table above.⁵ I use 9908 samples from the eight-dimensional parameter space. For each of these sets of parameters, I compute the discounted two-period profit from each of the four strategies. I then rank each of the strategies according to their discounted profit to determine which strategy the agent will choose. The following table computes the median value of each of the parameters conditional on a given strategy being optimal.

⁵The data and manipulations can be fully reproduced with the following Python Notebook: <https://colab.research.google.com/drive/1vQz0UHOiQwL6Ed7m9Q7mZeuVhMo8FB0-?usp=sharing>

Table 3: Median value of sampled parameters conditional on a given strategy being optimal

	$n(\bar{z})$	$n(\underline{z})$	A_0	β	δ	$I(n(\bar{z}))$	$I(n(\underline{z}))$	I_0
$\underline{z} \rightarrow \underline{z}$	7.0	5.9	44.0	0.4	0.5	80.4	52.6	46.9
$\bar{z} \rightarrow \underline{z}$	7.2	5.2	49.1	0.6	0.6	79.1	26.5	21.0
$\underline{z} \rightarrow \bar{z}$	6.9	4.5	38.2	0.4	0.3	82.9	42.6	61.4
$\bar{z} \rightarrow \bar{z}$	7.1	1.7	52.0	0.5	0.5	78.7	33.2	34.4

There are many important insights that we can draw from the table above. To do so, we can compare each of the values in the table to the unconditional median of that parameter. For example, the overall median of δ is approximately .5, but in the δ column of the $\underline{z} \rightarrow \bar{z}$ row, the median value of δ is 0.30, much lower than the overall median. This means that when the low quality, then high quality strategy is optimal, we expect there to be relatively low audience depreciation relative to when other strategies are optimal.

I will now briefly characterize the profile of parameters that produces each of the winning outcomes. First, if $\underline{z} \rightarrow \underline{z}$ is optimal, then there is probably, relatively little difference between the number of streams per capita of high and low quality art. Additionally, there is relatively little expected algorithmic benefit from high quality products. As $\underline{z} \rightarrow \underline{z}$ is the most pessimistic of the possible strategies, it is not surprising that this strategy arises when both the consumer and the algorithm are not very discerning.

For $\bar{z} \rightarrow \underline{z}$ to be optimal, then consumers should not be very picky, but the algorithm should. The algorithm should make it hard to new artists to gain exposure (small I_0). Additionally, the artist should have a stronger-than-average concern for the future $\beta = .6$, and a relatively high audience depreciation, $\delta = .6$. This description is consistent with intuition. If consumers are not very picky between high and low quality, but initial impressions are hard to come by, the artist will prioritize using more releases to grow their audience instead of relying on the algorithm.

The $\underline{z} \rightarrow \bar{z}$ is very rare among the points sampled, and it requires a unique mix of parameters to be optimal. Consumers should be moderately discerning, and the algorithm should punish low quality content, however, the algorithm also needs to favor new artists (I_0) more than low quality content $I(n(\underline{z}))$. Additionally greater present bias, and high audience retention benefit this strategy. This result, albeit rare, has important implications for how a streaming platform should treat its new artists. If the platform wishes to incentivize future high quality production from artists who start low quality, they should algorithmically favor new content.

The final case $\bar{z} \rightarrow \bar{z}$, is essentially the opposite of $\underline{z} \rightarrow \underline{z}$. For the high-quality only strategy to exist, both consumers and the algorithm should be very discerning of high quality content, those two conditions alone are usually sufficient to result in $\bar{z} \rightarrow \bar{z}$ as the optimal strategy as evidenced by the other variables being near their medians. This again has important implications for the streaming platform, if they wish

to establish an environment where high quality content thrives, they should ensure that their algorithm rewards high quality content, but the platform also relies on its consumers to be equally discerning.

Discussion of the table above that the discerningness of the consumer and the algorithm are important predictor of an artist's optimal behavior. As crude measures of discerningness, let $v = n(\bar{z})/n(z) - 1$ and $\iota = I(n(\bar{z})I/I(n(\bar{z})) - 1$. These variables are the percentage difference between the high and low quality versions of streams per capita and expected impressions per release. Using these two

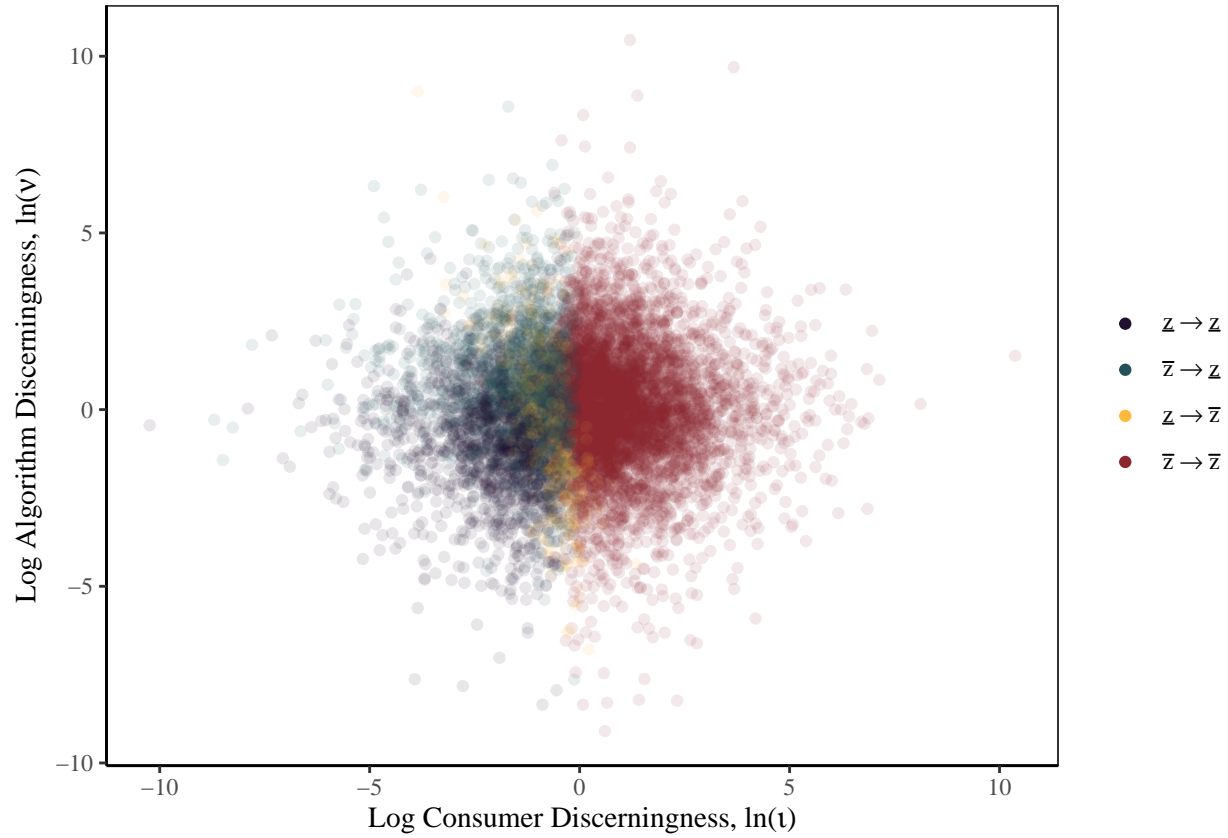


Figure 2: Consumer and Algorithmic Discerningness on Optimal Strategy

(I KNOW THE LEGEND IS MESSED UP THERE, BUT I'LL FIX IT LATER BECAUSE I WANT TO GET TO THE INSIGHTS)

Figure 2 (FIX LINK) shows that while there is some overlap between the optimal strategies, the regions are mostly distinct and contiguous. For example, whenever $\ln(v) > 0$ (high consumer discerningness, almost no other parameters matter and we are essentially guaranteed that the $\bar{z} \rightarrow \bar{z}$ is the ideal strategy. On the contrary when $\ln(v), \ln(\iota) < 0$ the ideal strategy is almost always $z \rightarrow z$. It's really quadrant II where $\ln(\iota) > 0$, but $\ln(v) < 0$, that the ideal strategy is least clear. Since high algorithmic discernment, and low consumer discernment push in opposite directions, it makes sense that quadrant II is ambiguous. What is less clear is why quadrant IV is not equally ambiguous. This seems to indicate that high consumer discernment outweighs low algorithmic discernment, but the inverse is not always true.

6 Next Steps, again

1. In that numerical simulation, see which strategy produces the biggest audiences for a given set of parameters
2. Do the sensitivity analysis of how shocks affect behavior
3. Get a better structure for the results section (though that may only come once I've narrowed down the results that are worth including).
4. Add the rest of the nocites not explicitly mentioned in the lit review

References

- Aguiar, L., & Waldfogel, J. (2018). Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music. *Journal of Political Economy*, 126(2), 492–524. <https://doi.org/10.1086/696229>
- Anderson, C. (2006). *The long tail: Why the future of business is selling less of more* (1st ed). Hyperion.
- Bender, M., Gal-Or, E., & Geylani, T. (2021). Attracting artists to music streaming platforms. *European Journal of Operational Research*, 290(3), 1083–1097. <https://doi.org/10.1016/j.ejor.2020.08.049>
- Elberse, A. (2008). Should You Invest in the Long Tail? *Harvard Business Review*, 11.
- Hiller, R. S., & Walter, J. M. (2017). The Rise of Streaming Music and Implications for Music Production. *Review of Network Economics*, 16(4), 351–385. <https://doi.org/10.1515/rne-2017-0064>
- Klein, B., & Leffler, K. (1981). The Role of Market Forces in Assuring Contractual Performance. *The Journal of Political Economy*.
- MacDonald, G. M. (1988). The Economics of Rising Stars. *The American Economic Review*, 78(1), 155–166.
- Rosen, S. (1981). The Economics of Superstars. *The American Economic Review*, 71(5), 845–858.
- Shapiro, C. (1983). Premiums for High Quality Products as Returns to Reputations. *The Quarterly Journal of Economics*, 98(4), 659–679. <https://doi.org/10.2307/1881782>
- Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), 355–374. <https://doi.org/10.2307/1882010>
- Thomes, T. P. (2013). An economic analysis of online streaming music services. *Information Economics and Policy*, 25(2), 81–91. <https://doi.org/10.1016/j.infoecopol.2013.04.001>
- Tom Scott. (2017). *Why The YouTube Algorithm Will Always Be A Mystery*. <https://www.youtube.com/watch?v=BSpAWkQLlgM>.
- YouTube stars Rhett and Link explain how algorithm changes supercharged their business to an estimated \$17.5 million in yearly income Indilens News Team ! Live Daily News On India and Around The World. (2020). In *Indilens ! Live Daily News on India*. <https://indilens.com/554159-youtube-stars-rhett-and-link-explain-how-algorithm-changes-supercharged-their-business-to-an-estimated-17-5-million-in-yearly-income/>.