Working title: Unequality of recommendation in the music industry

**Contents**

Abstract

Keywords

1. Introduction

2. Relationship between talent and popularity

2.1 Why and how does superstardom occur: economic theories of stardom

2.2 Empirical testing of stardom theories in the existing literature

3. Popularity gain via recommendation

3.1 Role of the social networks

3.2 Online content popularity and peer recommendations

3.3 Influence of the recommendation engines

4. Empirical testing on Last FM data

4.1 Methods

4.2 Model

4.3 Results

5. Conclusion

References

**2: Relationship between talent and popularity**

2.1: Why and how does superstardom occur: economic theories of stardom

As mentioned in the Introduction, we assume that there exists a popularity bias in recommendation systems. Therefore, it is valuable to understand what factors condition emergence of stars. Economists have been tackling this issue for the last 40 years. However, they have been often applying theories which had been formulated prior to their work.

Seventy years ago Marshall (1947) indicated that innovations in technology lower the unit price of quality goods and hence allow them to obtain greater market share. According to Rosen (1981), this very effect is partly responsible for the phenomenon of superstars. Exceptional revenues of stars are compelled by a market equilibrium that rewards people with increasing returns to ability. Due to scale economies allowing joint consumption, superstars can reach a vast audience – production costs do not rise in proportion to the size of seller’s market. In this way, a small number of suppliers can satisfy the demand of the whole market. Nonetheless, in large economies of scale, each member of the tiny group of artists can only reach high salary if the demand is highly concentrated on their services.

There are two substantial theories of stardom which reach back to the eighties. They are distinct but not mutually exclusive and state what drives the demand for superstar services: according to Rosen (1981) this factor is superior talent combined with perfect reproducibility of art, whilst Adler (1985) claimed network externalities of popularity to be responsible.

Pursuant to Rosen, poorer quality is an imperfect substitute for higher quality. He claims that small differences in talent translate to large earnings differentials. Then, most people are less satisfied with a performance of a less talented and cheaper artist when they have an opportunity to enjoy a top performance, even with the higher cost (Frey 1998). If the best artist is significantly better than the competition, “each consumer consuming the best” is a special case (Adler 2006). In this circumstances he becomes a monopolist whose profit maximizing strategy depends on the elasticity of demand for his product (if the demand is highly elastic, it pays off to serve the whole market). In Rosen’s model, there are two extreme options: either there is a very top artist who sets a high price and sells it to only a fraction of consumers (unless the demand is highly elastic) or there are several equally talented artists, one of which serves the whole market but is poor. In conclusion, if a star is both extraordinarily popular and rich, his talent must be unquestionably greater than the rest.

The second theory by Adler (1985) refers to the concept of “consumption capital” (Becker and Stigler 1977). Consumers built it in art and the larger it is, the greater the enjoyment from each encounter with its subject (art and artist). Consumption is a dynamic process rather than momentary experience and consumers want to consume the same art that others do, which is a key factor underlying production of superstars. When the artist is popular, it is easier to find other people familiar with his works or media coverage. However, in Adler’s model, the emergence of a star is a chance event: consumers first include artists randomly in their consumption basket, and it is “pure luck” that one artist ends up with more patrons than the rest. It obviously gives him advantage, which can then snowball into superstardom.

The third important theory of stardom comes from MacDonald (1988). He described a dynamic process through which stars arise. Each artist is capable of a good or bad performance. The difference in talent is then defined differently than in the two previous models: it is not the quality of performances but rather the probability that a particular performance will be good (constant throughout the artist’s career). But from the viewpoint of audiences, this probability is lower for a new performer than for a famous one. Those who perform poorly drop out, while good artists stay on the market and increase their probability of performing well in the future. In this way, artists with a good track record can appoint higher prices. In conclusion, artists of corresponding talent do equitably well.

Finally, the last important theory of stardom is a stochastic model of superstardom developed by Chung and Cox (1994). They employed a stochastic model of Yule (1924) and Simon (1995) (known as the “Yule distribution”) as a representation of the consumer’s choice behaviour. Simon showed that a wide range of data conforms to a class of distributions obtained from stochastic processes similar to those yielding negative binomial or log series distributions. This class is given by:

where *ψ* and *ρ* are constants, > 0, ∞ > ρ > 0 and:

2.2 Empirical testing of stardom theories in existing literature

In this section we deliver a brief summary of empirical literature on superstar emergence.

Hamlen (1991, 1994) measured relationship between harmonic quality of a singer’s voice and record sales and found out that they indeed increase with the quality of the voice. However, in this studies, differences in talent exceeded differences in sales. The author interpreted results being conforming with Rosen’s model (artists are rewarded for talent). The above mentioned article of Chung and Cox also contains empirical part. They proved that the distribution of success among artists resembles the Yule distribution basing on the number of CDs sold. The probability that a consumer would buy a particular CD increased with the number of previous sales of that CD. However, a small chance that a consumer will choose a new CD that no other person has bought always remains. Hence, an initial small advantage may snowball into success, which supports Adler’s theory. Schulze (2003) argues with both above mentioned works. He criticized Hamlen for the quality of voice being irrelevant measure for singles of non-classical genres. He also claims that the process described by Chung and Cox is also consistent with Rosen’s theory (consumers’ choices based on talent).

Another study advocating importance of initial advantage is the one by Ginsburgh and van Ours (2003) who investigated indicators of success in the Queen Elizabeth Piano Competition. They showed that the randomly assigned order in which competitors perform influences the result of competition. It means that random success in the competition affects success on market. A possible explanation of this phenomenon is that success in the competition serves as a focal point for conformist consumers rather than a token of talent.

There is yet another different field of study, in which determinants of success are measured, namely sports. Its industry is believed to be related enough to display patterns similar to those in art. Franck and Nuesch (2008) check the influence of on-field performance and media publicity on the emergence of soccer superstars in Germany. They used cross-sectional samples and assumed talent indicators exogenous proving that both talent (measure by professional expertise) and popularity increase demand for sport stars. Lehmann and Schulze (2008) adopted a similar approach and regressed salary proxies of 359 players on three performance measures and number of citations in an online magazine. They found evidence contrary to Franck and Nuesch: neither performance nor publicity explained salaries at the 5% significance level. Two years later those authors performed a similar study (Franck and Nuesch 2010) but allowed for the correction of teams. In that case, a single player’s talent was considered his contribution to the team output. A team production function was therefore estimated to detect critical elements that affect a team’s success. They found evidence that both talent and non-performance-related popularity contribute to the market value differentials in German soccer league.

When adapting results of stardom studies in sports to art industry, one must bear in mind minor differences between those fields. The most important one is the competitive nature of sport. In a given competition, every sportsman must accomplish the same task. On the other hand, artists have more opportunities to express themselves and display talent without artificial limitations. Therefore, in art, unlike sports, there are no measurable standards. Moreover, talent in art is much harder to define. For example, some music appeals to a subset of listeners but not others, because the quality of arts is highly subjective. Nonetheless, both Rosen (1981) and Adler (1985) assume a homogenous consumer. They argue that heterogeneity of tastes does not deteriorate the mechanisms of superstar emergence but confines a producer’s market, which means that consumers of similar preference comprise a market with characteristic stars.

The relationship between talent and popularity has been also measured in an experimental way. Salganik, Dodds and Watts (2006) conducted an experimental study in an artificial cultural market. 14,341 participants downloaded new songs either (the order assigned randomly) with or without knowledge of the previous participants’ choices. It turned out, that knowing choices of other “customers” contributed to inequality and unpredictability of the market. Social influence enhances the skewness of the market distribution and uncertainty of success. However, in this study the outcome yielded capricious even in a set up without knowledge about download statistics. The authors concluded that no metric of a record’s quality can accurately forecast success.

As shown in this section, the literature on the relationship between talent and stardom is practically limited to a few fundamental theories. They are decidedly worth testing. However, many different approaches remain to be applied.

**3: Popularity gain via recommendation**

3.1 Role of the social networks

Previously mentioned theories on stardom try to explain the relationship between market factors and an individual’s popularity. However, in the optics of growing field of social network analysis, there is a growing tendency of including network effects and relationships between market actors in the research of popularity.

Social networks are often represented by mathematical models, aligned with graph theorem. Each participant is node and his/her connection with other person is called an edge. Originally, social interactions were studied in real life conditions. Analysis of those is concentrated on relationships between individuals, that are constructed via socialising (Haythornthwaite, 2005). Social structures build by actors affect their behaviour (Doreian, 1989). This influence between participants was proven in many fields. One of the most famous study of social networks effects was a paper by Christakis and Fowler (2007) that connected social structures with obesity spread. That paper started a large scaled discussion on social networks, but also raised some critique – mainly connected to the exclusion of environmental factors in that particular analysis (Haythornthwaite & Fletcher, 2008). Nonetheless, in the overall literature there is no doubt that social structures influence people’s beliefs. Basing on this observation, there exists a large field of studies that uncovers connection of social networks and popularity – regardless of the fame’s object, which can be an idea, past time activity, fashion trend or an artist (Ellison et al., 2009; Centola, 2010; Valenzuela, et al., 2012; Yamaguchi, et al., 2014). In those cases researchers stress that network participants are more likely to be interested in an issue popular amongst other egos. Spreading an idea or a certain lifestyle happens more freely in a group of previously connected individuals (Lind, 2007).

Modern studies concentrate more on the virtual realisations of people’s interactions. Online websites and social media are the newest platform that allows for the more dynamic information flow and accelerate new social interactions (Galuszka, 2015). There are research suggesting that, through the online media, an individual can have and maintain more social contacts that in the real life (Chou & Peng, 2007). By nature, online relationship are more shallow than the latter and are consisted mainly around a common interest or a hobby (Ren et al., 2007). However this type of interaction is of key importance for online attractiveness scores.

3.2. Online content popularity and peer recommendations

Popularity of online content is highly influenced by social networks (Jamali & Rangwala, 2009). In this case real life connections are not as important, as the activity of followed accounts. In online environment actions of other users are often a guideline for an actor’s individual behaviour (Smith et al., 2005). This copying mechanism leads to the boost of popularity of other agents’ liked content. If one person is followed (observed and often copied) by many other users (in network terminology: its node leads to many edges), their recommendation for certain product can lead to a rise in its sales (so called *influencer marketing*) (Brown & Hayes, 2008). In a case of egos with less edges the quantity counts – a advice given by a large number of users will most probably lead to the change of behaviour of an individual agent (Smith et al., 2005).

Regarding online social networks, many recommendation mechanisms are happening inside them. Social groups, and networks structured within them, are often consisting of people with similar attributes (Hui & Buchegger, 2009). One type of the similarity can be a similar music taste or even belonging to the same fandom (Galuszka, 2015). In a group connected by a mutual trust and analogous preferences a suggestion given in a field of common interests is worth more than in a group of strangers. Existing social network strengthens the recommendation and grants it more influence over the agents (Hanna et al., 2011).

In a music industry, a user can often observe representation of networks in which their liked artist is active. In this case it can be expressed by cooccurrences in songs, guest performances on tours or photos shared with different artists on social media (Morris, 2014). Those signals often serve marketing purposes – cooccurrences boost artist’s popularity by promoting them among other performer’s fans (Fields, 2008). There are also many examples of the usage of existing social networks and celebrity of already successful artist, by aspiring performers – some of the strategies are visiting popular music festivals with alike musicians in a line-up or performing as supports before the main artist for a few concerts or a whole tour (Leenders, 2005).

Social connections, or lack of thereof, often determine the popularity of a certain idea. In the music industry, where the revenue is strongly correlated with the fame, any factor that can improve the chances for higher popularity should be taken into consideration by a rising star.

3.3 Influence of the recommendation engines

Individual talent and charisma are accounting only for a part of a star’s success. In the modern music industry, dominated by streaming platforms, the access to the new listeners is a key to the new-comer’s achievement – and for this the recommendation engines are one of the most important factors (Datta et al., 2017).

Streaming platforms allow the user to play music on demand for a small monthly fee or even without it. Nowadays, many of them offer tailor-made playlists for a specific user, that help to discover new songs and artists that suit the music taste of an agent (Prey, 2018). Those are the results of recommendation engines work. Their job is to analyse user’s behaviour and music choices and match them with fresh, undiscovered tracks (Pichl et al., 2016).

Most of them are pre-trained machine learning algorithms, that are based on three main aspects: collaborative filtering, natural language processing (NLP) and audio assessment models (Chen & Chen, 2001; Ciocca, 2017). First one utilizes the information about users’ choices – using streaming history of each agent, the model seeks for similarity. When a high similarity between a group users behaviour is found, the engine is likely to recommend the other users’ choices, that have not been discovered by the individual yet (Cohen & Fan, 2000). The second filter is based on a text analysis – it may include analysis of lyrics or even online websites tracking for cooccurrences of artists in one article or similarity of adjectives used to describe them in different sources (Wang et al., 2013; Hyung et al., 2014). The last one is based on audio models, that look for musical resemblance in songs and recommend tracks that sound familiar to the user (Schnitzer, 2012; Ciocca, 2017).

There are many concerns regarding the engines’ activity. In their nature, they are only algorithms pre-trained for the recommendation purposes, mostly consisting of different layers of neural networks, which specific parts cannot be easily explained to the observer. Their influence over music industry is more than visible – there are many cases in which a musician gained popularity solely by streaming platforms (Page & Ning, 2014). Their role is believed to be crucial in determining success or failure of a particular artist.

A large number of concerns is connected to possible biases of those systems. Algorithms are often described as inequality rising, mainly because of collaborative filtering usage (Yao & Huang, 2017). They are advocating artists that are already popular among other users. This phenomenon increases the popularity of previously famous individuals and it is often called the ‘popularity bias’ (Celma & Cano, 2008). The highly streamed artists are recommended not for their talent, but solely for their current popularity (Abdollahpouri et al., 2019). The NLP layer of engines can also be biased toward already recognised artists. References in articles are often linking performers to their popular equivalents, boosting the fame of the latter. It is worth noting, that only the audio models are free from this bias, as they take into account no more than the musical layer of a song, without any popularity undertones or social connotations.

The unfairness of recommendation engines is a sound ground for many research in a field (Kowalke, 2015; Yao & Huang, 2017; Abdollahpouri et al., 2019). However, there also many supporters than emphasize their role in enhancing diversity and giving a chance for a less recognised artists to grow (Zhang & Hurley, 2008).

There is also another issue concerning the recommendations. The systems are utilizing a basic bias that is attributed to human cognition – *we like what we know* (Bohrn et al., 2013). This means, that people more often want to enjoy art that are acknowledged with, that discover entirely different genres they may eventually like (Leder, 2001). The desire for the similar and known is so strong, that even without recommendation engines, biased or not, users will end up listening their old, favourite artists and songs.

**References**

Marshall, A. Principles of Economics. 8th ed. New York: MacMillan, 1947.

Rosen, S. “The Economics of Superstars.” American Economic Review, 71(5), 1981, 845–58.

Adler, M. “Stardom and Talent.” American Economic Review, 75(1), 1985, 208–12.

Frey, B. S. “Superstar Museums. An Economic Analysis.” Journal of Cultural Economics, 22, 1998, 113–25.

Adler, M. “Stardom and Talent” Handbook of the Economics of Art and Culture, Volume 1 41, Chapter 25, 2006 Elsevier B.V.

Stigler, G., Becker, G. (1977). “De gustibus non est disputandum”. American Economic Review 67, 76–90.

MacDonald, G. (1988). “The economics of rising stars”. American Economic Review 78, 155–166.

Chung, K., Cox, R. (1994). “A stochastic model of superstardom: An application of the Yule distribution”. 6 Review of Economics and Statistics 76, 771–775.

Simon, Herbert A., "On a Class of Skew Distribution Functions," Biometrika 42:3/4 (1955), 425^40.

Yule, G. Udny, "A Mathematical Theory of Evolution, based on the Conclusions of Dr. J. C. Willis, F. R. S.," Philosophical Transactions of the Royal Society of London, Series B, 213 (1924), 21-87.

Hamlen, W. (1991). “Superstardom in popular music: Empirical evidence”. Review of Economics and Statistics 73, 729–733. 11

Hamlen, W. (1994). “Variety and superstardom in popular music”. Economic Inquiry 32, 395–406.

Schulze, G.G. (2003). “Superstars”. In: Towse, R. (Ed.), Handbook of Cultural Economics. Edward Elgar, Cheltenham, pp. 431–436.

Ginsburgh, V., van Ours, J. (2003). “Expert opinion and compensation: Evidence from a musical competition”. American Economic Review 93, 289–296.

Franck, E., and S. Nuesch. “Mechanisms of Superstar Formation in German Soccer: Empirical Evidence.” European Sport Management Quarterly, 8(2), 2008, 145 – 64.

Lehmann, E., and G. G. Schulze. “What Does it Take to Be a Star? —The Role of Performance and the Media for German Soccer Players.” Applied Economics Quar- terly, 54(1), 2008, 59–70.

Franck, E., and S. Nuesch. TALENT AND/OR POPULARITY: WHAT DOES IT TAKE TO BE A SUPERSTAR? Economic Inquiry Vol. 50, No. 1, January 2012, 202–216

Salganik, M. J., P. S. Dodds, and D. J. Watts. “Experimen- tal Study of Inequality and Unpredictability in an Artificial Cultural Market.” Science, 311(5762), 2006,

854 – 56.

Celma, Ò., & Cano, P. (2008, August). From hits to niches?: or how popular artists can bias music recommendation and discovery. In *Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition* (p. 5). ACM.

Doreian, P., Freeman, L., White, D., & Romney, A. (1989). Models of network effects on social actors. *Research methods in social network analysis*, 295-317.

Cohen-Cole, E., & Fletcher, J. M. (2008). Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of health economics*, *27*(5), 1382-1387.

Haythornthwaite, C. (2005). Social networks and Internet connectivity effects. *Information, Community & Society*, *8*(2), 125-147.

Smith, D., Menon, S., & Sivakumar, K. (2005). Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of interactive marketing*, *19*(3), 15-37.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine*, *357*(4), 370-379.

Galuszka, P. (2015). New economy of fandom. *Popular Music and Society*, *38*(1), 25-43.

Jamali, S., & Rangwala, H. (2009, November). Digging digg: Comment mining, popularity prediction, and social network analysis. In *2009 International Conference on Web Information Systems and Mining* (pp. 32-38). IEEE.

Yamaguchi, K., Berg, T. L., & Ortiz, L. E. (2014, November). Chic or social: Visual popularity analysis in online fashion networks. In *Proceedings of the 22nd ACM international conference on Multimedia* (pp. 773-776). ACM.

Valenzuela, S., Arriagada, A., & Scherman, A. (2012). The social media basis of youth protest behavior: The case of Chile. *Journal of Communication*, *62*(2), 299-314.

Hui, P., & Buchegger, S. (2009, July). Groupthink and peer pressure: Social influence in online social network groups. In *2009 International Conference on Advances in Social Network Analysis and Mining* (pp. 53-59). IEEE.

Ellison, N. B., Lampe, C., & Steinfield, C. (2009). Social network sites and society: Current trends and future possibilities. *interactions*, *16*(1), 6.

Centola, D. (2010). The spread of behavior in an online social network experiment. *science*, *329*(5996), 1194-1197.

Chou, C., & Peng, H. (2007). Net-friends: Adolescents’ attitudes and experiences vs. teachers’ concerns. *Computers in Human Behavior*, *23*(5), 2394-2413.

Ren, Y., Kraut, R., & Kiesler, S. (2007). Applying common identity and bond theory to design of online communities. *Organization studies*, *28*(3), 377-408.

Page, W., & Ning, K. (2014). Anatomy of a hit: How Meghan Trainor made the UK chart without selling downloads. *Spotify Insights*.

Ciocca, S. (2017). Spotify’s Discover Weekly: How machine learning finds your new music. *Hackernoon. URL: https://hackernoon. com/spotifys-discover-weekly-how-machine-learning-finds-yournew-music-19a41ab76efe*.

Chen, H. C., & Chen, A. L. (2001, October). A music recommendation system based on music data grouping and user interests. In *Proceedings of the tenth international conference on Information and knowledge management* (pp. 231-238). ACM.

Cohen, W. W., & Fan, W. (2000). Web-collaborative filtering: Recommending music by crawling the web. *Computer Networks*, *33*(1-6), 685-698.

Hyung, Z., Lee, K., & Lee, K. (2014). Music recommendation using text analysis on song requests to radio stations. *Expert Systems with Applications*, *41*(5), 2608-2618.

Wang, M., Kawamura, T., Sei, Y., Nakagawa, H., Tahara, Y., & Ohsuga, A. (2013, November). Context-aware music recommendation with serendipity using semantic relations. In *Joint International Semantic Technology Conference* (pp. 17-32). Springer, Cham.

Schnitzer, D., Flexer, A., & Widmer, G. (2012). A fast audio similarity retrieval method for millions of music tracks. *Multimedia Tools and Applications*, *58*(1), 23-40.

Kowalke, C. (2015). How Spotify Killed the Radio Star: An Analysis on How the Songwriter Equity Act Could Aid the Current Online Music Distribution Market in Failing Artists. *Cybaris Intell. Prop. L. Rev.*, *6*, 193.

Yao, S., & Huang, B. (2017). Beyond parity: Fairness objectives for collaborative filtering. In *Advances in Neural Information Processing Systems* (pp. 2921-2930).

Abdollahpouri, H., Mansoury, M., Burke, R., & Mobasher, B. (2019). The unfairness of popularity bias in recommendation. *arXiv preprint arXiv:1907.13286*.

Zhang, M., & Hurley, N. (2008, October). Avoiding monotony: improving the diversity of recommendation lists. In *Proceedings of the 2008 ACM conference on Recommender systems* (pp. 123-130). ACM.

Bohrn, I. C., Altmann, U., Lubrich, O., Menninghaus, W., & Jacobs, A. M. (2013). When we like what we know–A parametric fMRI analysis of beauty and familiarity. *Brain and Language*, *124*(1), 1-8.

Leder, H. (2001). Determinants of preference: When do we like what we know?. *Empirical Studies of the Arts*, *19*(2), 201-211.

Brown, D., & Hayes, N. (2008). *Influencer marketing*. Routledge.

Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We’re all connected: The power of the social media ecosystem. Business horizons, 54(3), 265-273.

Fields, B., Casey, M. A., Jacobson, K., & Sandler, M. B. (2008, August). Do You sound like your Friends? Exploring Artist Similarity via Artist Social Network Relationships and audio signal Processing. In *ICMC*.

Morris, J. W. (2014). Artists as entrepreneurs, fans as workers. *Popular Music and society*, *37*(3), 273-290.

Leenders, M. A., van Telgen, J., Gemser, G., & Van der Wurff, R. (2005). Success in the Dutch music festival market: the role of format and content. *International Journal on Media Management*, *7*(3-4), 148-157.

Datta, H., Knox, G., & Bronnenberg, B. J. (2017). Changing their tune: How consumers’ adoption of online streaming affects music consumption and discovery. *Marketing Science*, *37*(1), 5-21.

Pichl, M., Zangerle, E., & Specht, G. (2016, December). Understanding playlist creation on music streaming platforms. In *2016 IEEE International Symposium on Multimedia (ISM)* (pp. 475-480). IEEE.

Prey, R. (2018). Nothing personal: algorithmic individuation on music streaming platforms. *Media, Culture & Society*, *40*(7), 1086-1100.

Lind, P. G., da Silva, L. R., Andrade Jr, J. S., & Herrmann, H. J. (2007). Spreading gossip in social networks. *Physical Review E*, *76*(3), 036117.