The impact of recommender systems on competition between music companies

Consultation response to the CMA's Music and streaming market study*

Peter L Ormosi†
Norwich Business School, and
Centre for Competition Policy
University of East Anglia
NR4 7TJ, Norwich, UK
[redacted]

Rahul Savani
Department of Computer Science
The University of Liverpool
L69 7ZX, Liverpool, UK
[redacted]

February 16, 2022

Autonomous recommender systems (RS) have become a key component of music streaming. These systems are responsible for composing popular playlists, or selecting the songs that follow the song that a listener has just finished listening. One of the reasons RS have become so ubiquitous on music streaming platforms is due to the fact that online listener behaviour is often characterised by cognitive biases: people prefer taking mental shortcuts, rather than tediously evaluating the endless number of music consumption choices they face. RS facilitate these mental shortcuts by recommending content that best satisfies the listeners' needs.

RS in general can make for more efficient outcomes, and there is a sizeable body of evidence of these benefits from other markets (Zhang 2018, Waldfogel 2017, Zentner et al. 2013, Brynjolfsson et al. 2011). On the other hand, the rigid architecture of a RS, and the inherent biases in their design, amplified by continuous feedback loops between the users and the RS, can limit free exchange and lead to inefficient outcomes. Moreover, if the RS itself produces a biased recommendation, it will be further amplified by the platform users' cognitive biases. The purpose of our response to the CMA's consultation, is to draw attention to how these biases can distort competition on the supplier side of the streaming platform's market (music companies), create significant entry barriers for some suppliers whilst favouring others, homogenise taste and choices, and disincentivise innovation.

Our response draws attention to the potentially anti-competitive effects of biases due to the design of the RS pipeline (from data collection, through the design of the recommender

*This response draws from Fletcher, Ormosi, and Savani (2022) Recommender systems, and their impact on market competition, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4036813.

†Corresponding author.

model, to the choice architecture design). Therefore our focus is not on the overtly or tacitly anti-competitive conduct by the platform or the music companies, rather the areas that could constrain free competition even in the absence of malintent the platform. That said, the biases discussed in this response can be easily abused by platforms to further harm competition and consequently consumers.

Below, for notational simplicity, we will use the shorthand expression *platform* for the streaming platform, where listeners (denoted as *users* below) access content provided by music companies or music creators (we refer to these as *suppliers*). In this response we do not offer an introduction to how recommender systems work, this can be found in our working paper.

1 Biases in the RS pipeline

Most music recommenders work by looping over a number of consecutive steps. To offer a stylised, three-step example, first, data is collected from the users and on the items available on the platform. Second, the data is organised and entered into a recommender model. Finally, the recommendation is fed back to the user. Figure 1 offers a visual summary of this pipeline.

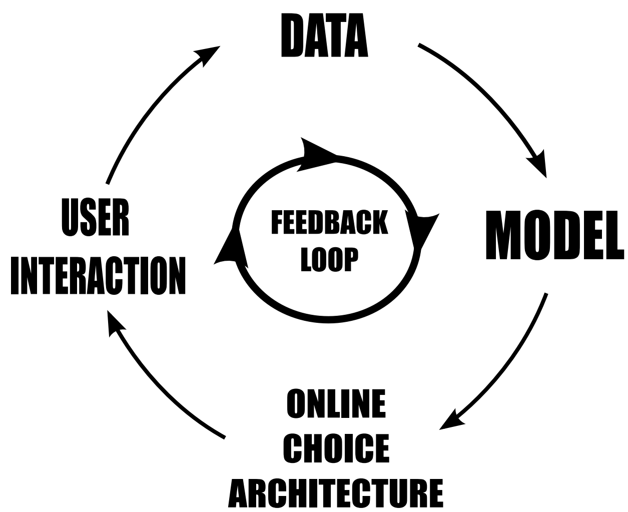

Figure 1: The role of feedback loops in the RS pipeline

1.1 Biases in the data used for the RS

The data collection process can be the source of some of the biases in RS. For example, in explicit data, the reviews/ratings given to items on a platform are not randomly distributed. Users are more likely to review songs that they associate with more extreme feedback (very bad or very good). This **selection bias** could mean the some items are less likely to be reviewed, creating a biased dataset where the recommender has to learn from many reviews given to a few items and no explicit information on the others. Similarly, users are more likely to provide feedback on the few items they are most likely to be exposed to (exposure could be due to popularity bias, but also due to positioning bias, i.e. how the given item is positioned in front of the user). If the RS is unable to distinguish between no-feedback because of no interest, and no feedback because of no exposure, it can make this type of exposure bias very difficult to de-bias.

There can also be serious economies of scale and scope related to data. This means that the markets for RS tend themselves to exhibit incumbency bias (due to data feedback loops) and to be highly concentrated (due to the economies of scale and scope in data).

1.2 Bias from the recommender methods

Probably the most obvious bias related to the RS design is **popularity bias**, i.e. more popular content or suppliers are more likely to be recommended, even if a less popular one would be more closely aligned with the consumer's preferences. Collaborative filtering based recommenders (which rely on the assumption that users with a similar listening history have similar consumption preferences) have a tendency to recommend the most popular content, i.e. items with most user engagement data. Popularity bias can stem from simple data availability issues. Popular content receives more listening, more ratings, and more consumer interaction. All of this largely enhances the data available to be used in the RS. On the other hand, less popular or new items on the platform have fewer (if any) user interaction, which may feed into either (1) less likely recommendation, or (2) lower quality recommendation, which means lower user satisfaction. Before we move on, it is important to add that popularity bias may be simply an indication that some content are superior to others. It is not obviously clear how to distinguish between desirable and undesirable sources of popularity bias, but it is something that has been looked at by some researchers (Ciampaglia et al. 2018, Zhao et al. 2021).

Recommender systems disproportionately favour incumbent products/suppliers - for which they have data - over new entrant. The literature knows this as the 'cold start' problem or **incumbency bias**. This can be caused simply by the fact the there may only be limited data available for new products/entrants. But it can also stem from the RS design. For example under collaborative filtering, if no user has ever listened to a new song, it is unlikely to be recommended to other similar users. Content based filters can help mitigate this issue.

Content based filter look at the features of the songs to find similarity and then link this to data on the songs the user engages with. Content based RS also disproportionately favour individual consumers items that are 'similar' to what they have consumed before (**individual homogeneity bias**). As content-based filtering methods rely on recommending content similar to what a user has already consumed, they are likely to create filter bubbles, where each individual faces a more homogenous choice. This leaves very limited scope for a surprise factor. A related issue is cultural homogeneity bias, which can stem from situations where the recommender disproportionately favours songs that fit with cultural norms or norms for particular identified groups (for example favouring English language music) over more niche items (probably due to weak RS design).

1.3 Bias in the choice architecture

Platforms can enhance and influence consumer's purchasing decisions by designing the choice architecture on the user interface. In practice, recommender systems, by definition, always involve some choice architecture. For example, a streaming platform can help consumer decision-making by bringing together choices in one place and reducing search costs. But in presenting these, they necessarily have to choose which songs or albums to present and in what order, or offer some choices more prominence than others. This inherently creates a choice architecture, which in turn will tend to affect consumers' listening behaviour.

One might argue that if the streaming platform is neutral, these nudges should not create any bias. But even if the platform does not strategically differentiate between suppliers, the biases inherited from the previous steps of the RS pipeline (data collection and modelling), combined with digital nudges, are more likely to lead to enhance those previous biases, and reinforce content homogenisation through the choice architecture.

The sequencing of songs, (another form of online choice architecture) is a good example of the confounding effect of bias inherited from an earlier stage of the RS pipeline. Sequencing also includes automated playlist continuation (or next song recommendation), for example where a listener's chosen playlist has finished playing. If the RS suffers from popularity bias, it is likely to also manifest in which song is recommended next, and popular ones are most likely to come first (Vall et al. 2019). As the probability of being listened to diminishes the further down a song is on a playlist, this can further amplify the popularity bias.

1.4 Feedback loops

As the RS finishes one iteration of the recommendation pipeline (data collection, learning of model, feedback to user), and the recommendation is returned to the user, the user will decide whether to engage with the recommendation or ignore it. This behaviour then becomes new data that is fed back into the learning model, and the new recommendation is returned to the user again. These loops then continue as long as the user is engaged with the platform.

If there is bias at any stage in the loop, it will be reinforced and further exacerbated through subsequent loops. Take the example where the RS recommends a popular song instead of the song that would best fit with the user's actual or potential preferences, simply because there is more data available on the popular song, or because the RS model put too much weight on collaborative filtering. As the user is then more likely to engage with the popular song, this information then feeds back into the RS loop and further amplifies the popularity bias.

As the biases are amplified, diversity falls, and it eventually reinforces the homogenisation of choice ('filter bubbles' or 'echo chambers') (Mansoury et al. 2020). Algorithmic confounding can make this effect more severe, i.e. when the human behaviour that we measure is already exposed to algorithmic recommendations (Chaney et al. 2018). For example, on a recommended song is based on the listening patterns of other users who themselves had already been affected by the RS.

2 The supplier side impact of RS biases

The RS deployed on streaming platforms inevitably affect the supplier side of the market. If the RS produces biased recommendations, these biases will make their effect felt on the competition between music companies. Below we offer a few examples how the above identified biases can affect competition between suppliers. It is important to emphasise that the concerns raised below are not contingent on malicious intent by the streaming platform to limit competition. Instead, our focus is on the anti-competitive effects induced by the characteristics of how some recommender methods work as introduced above.

2.1 Entry barriers created by bias in the RS

As popular or incumbent music features in more recommendations (popularity and incumbency bias), the chance that more users engage with these recommendations increases. In this setting new entrants have very little chance of being recommended. The problem with these biases is an inherently dynamic one especially when it comes to evaluating its impact on competition between music companies. A streaming platform with a RS that suffers from popularity or incumbency biases will shun opportunities for new entrants, and is bound to converge to a concentrated one with a few large suppliers. Moreover, it is the ability of streaming services to give more prominence to long-tail products (especially when compared to traditional retail channels) that popularity bias risks losing. Access to the long-tail is a key driver of growth, and any deterioration to it would harm the economy and consumers alike.

The choice of the recommender method may seem like a technical question, but it is one that can directly contribute to these entry barriers. Collaborative filtering struggles with the problem of recommending new items given its reliance on information from previous purchases of the same item. Songs with more listening history data (i.e. more popular songs) are more likely to be recommended simply because they feed more data into the RS (cold start problem). For example, Spotify estimates that around 20% of their songs are never listened to at all. A RS with too much weight on content-based filtering can also lead to entry barriers to new products by creating filter bubbles, or narrow 'echo chambers' that become more extreme in their outlook due to feedback loops.

Popularity bias is perpetuated by the continuous feedback loops from listeners' engagement with the recommendation. As the RS has a tendency to recommend the most popular music, it acts as an entry barrier to new (potentially innovative content) and reduce the present value of such innovative content (because revenue on streaming platforms is typically linked to streaming numbers), which in turn could dampen innovation and the creation of new content. Biases in the RS lead to a fall in the diversity of products that each customer is recommended. Although at an aggregate level there may be a more diverse range of content recommended, each listener is aware of a more limited range of content (Holtz et al. 2020). In the long run this can lead to increased homogeneity in cultural content.

2.2 Increased homogeneity and increased segmentation

'Popularity bias' tends to reduce sales of items in the long tail. This will tend to reduce variety and increase homogeneity. This is more serious than it may sound, because the growth of the long-tail sales has been one of the successes of the digital economy to date and a key driver of growth (Brynjolfsson et al. 2011). Any deterioration of its viability would harm the economy and consumers alike. In addition, if RS disproportionately recommend similar items to those that a consumer has already purchased ('homogeneity bias'), this can potentially also have a market-wide effect, increasing product homogeneity and reducing variety.

It should be noted that increased homogeneity can appear to be good for competition, as consumers face more homogeneous sets of choices, and prices tend to be lower in homogeneous product markets (**?**). However, even if the short term effect of homogeneity bias leads to more price competition (at the individual level) between the most popular items, it can also lead to reducing competition in other dimensions such as range, quality, service, innovation or general market dynamism. If innovative new products cannot expect to gain sales, the R&D required will not be funded.

If specific categories of consumers are given disproportionately similar recommendations ('conformity bias'), this can lead to 'filter bubbles' or 'echo chambers'. While concerns in this area tend to focus on the cultural implications, there are also competition risks arising. Products can become identified with one social group, and therefore never recommended to consumers in another social group

On the positive side, such market segmentation can potentially act to increase market-wide variety, even if each individual group receives more homogeneous recommendations (Holtz et al. 2020). However, it can also reduce competition within each social group, since from a particular social group's perspective item x may not viewed as a substitute for item y, even if they effectively fulfil the same function.

2.3 Tilting the RS to introduce platform interest

Streaming platforms receive revenue from subscription fees or from advertising. The platform is interested in maximising this revenue, but also in sustaining high user-satisfaction. A number of previous works have shown that a recommender system that maximises revenue, does not necessarily harm other objectives, such as user satisfaction (Azaria et al. 2013). This is in line with intuition. Maximising user satisfaction is likely to be directly proportional to user engagement, which is pivotal for all sources of platform revenue (more satisfied users are likely to spend more time on the platform, or more user engagement converts to more advertising revenue). But what happens if there are two songs, offered by supplier A and B respectively, and recommending A would result in a higher level of user satisfaction, but recommending (and the user listening to) B represents higher revenue for the platform. Bourreau & Gaudin (2018) suggests that in these cases (i.e. where two products on the platform are associated with different revenues for the platform) it is possible that product B would be recommended.

In this response so far, we have assumed that the intention of RS design is user-centric (even if this intention is not always achieved). In practice, however, platforms are commercial enterprises. As such, their primary interest is typically in generating profit for their shareholders. As explained above, in relation to RS design the interest of shareholders are, to a large extent, aligned with those of consumers. But it is possible that the RS's profit maximising interest goes against being user-centric.

The fact that recommenders nudge the users towards not selecting the content that would maximise their utility can impact consumer welfare. Zhang et al. (2021) estimate this effect through a field experiment on video recommenders, and find that when the recommender maximises profit (or revenue) rather than consumer surplus by strategically placing songs or albums in salient slots, can harm consumer welfare. What makes this really difficult to assess is that deviating from the user's preferences to offer something surprising, in the expectation that the user will like it (a welfare enhancing case), may be difficult to distinguish from the case where the RS deviates from the user's preferences to pursue a different objective, such as revenue maximisation (a potentially welfare reducing case). Small mis-recommendations may not even be detected by the user, but given enough iterations of the recommendation loop, each of these small deviations can add up to a large sway away from what the user really wants.

There is also the possibility that the platform is vertically linked to music companies, whose own products compete with third party suppliers on the platform (for example Warner's ownership in Deezer, or Sony's and Universal's ownership in Spotify). In this case, the platform may prefer recommending B (self-preferencing), even if the user would have a preference for A. The details of course hinge on the impact of the recommendation, both on the marginal revenue from user satisfaction and on the marginal revenue from other sources. If the latter dominates, the user will not be recommended the song that is best aligned with her preferences. The market power of the platform is likely to affect how much weight each platform can attribute to maximising its revenue/profit to the detriment of maximising user satisfaction. Self-preferencing and preferential treatment of selected suppliers is not a novel idea. Platforms have been shown to favour their own products. What makes this different, is the subtlety of how the same can be achieved through RS. It is enough if the platform only marginally tips the platform to favour some products - feedback loops can amplify the impact of even the smallest changes to the initial conditions of the RS. With enough iterations through the RS loop, these small changes can completely alter the fate of some products and suppliers on the platform. For example, a self-preferencing platform may rank their own products above others for a short period and then stop. Feedback loops mean that this initial push will be preserved and amplified in the RS, and it may be enough to completely tilt the competitive playing field on the platform. Proving any misconduct - especially where the misconduct may not even be discernible from the genuine operation of the RS - seems impossible in these cases.

Moreover, given the bargaining power of major record labels (through the enormous catalogues of music they hold), it is also possible that they can use this position to influence the streaming platform. A relatively little-discussed implication of network effects is that certain suppliers can effectively become "must have". Without their presence on a given platform, consumers on the other side of that platform would switch to an alternative platform. This could in turn lead to other suppliers leaving, and so on. Such critical suppliers have substantial bargaining power and can potentially utilise this to require preferential treatment by a platform's RS. This requirement can be direct, but it can also be indirect. For example, some music streaming services have minimum payment guarantees with the three major record labels, each of which has substantial bargaining power. At the margin, such minimum payment guarantees may be expected to incentivise the streaming services to favour major label music over independent music (**??**).

A platform may also have wider strategic reasons for distorting supplier competition. For example, a firm which offers an ecosystem with many different services within it may wish to keep consumers within its 'walled garden'. As such, even if it does not itself provide a particular product, it may be more inclined to recommend a third-party product that lies within the walled garden than one which would take consumers outside it. For example, Google's mobile search service tended (until recently) to recommend content which was cached on Google's own AMP servers (AMP originally stood for 'Accelerated Mobile Pages'). This may be – as Google claimed – because Google could be sure of the download speed and quality of such content. However it might also have reflected Google's preference to keep consumers within the Google ecosystem. Similar considerations apply in relation to Amazon giving preference in its rankings to third party suppliers that use its 'fulfilled by Amazon' service.[1]

Another potential issue is to do with the data used for the RS and of the RS design used by the platform. This could be valuable for suppliers, as it can be used to predict which products are most likely to be recommended to consumers. If the platform is the only one with access to this information, it may help them create products that the consumers would more likely buy than those produced by third party suppliers. For example, in creating House of Cards, Netflix relied heavily on recommender data in its design, development, and talent selection (Schrage 2020, p. 11).

Transparency, in the objectives followed by the RS can be immensely helpful to make this distinction. On most (if not all) platforms the user is not aware whether she receives the recommendation because this truly reflects her personal preferences, or because it is what maximises the revenue of the platform. As long as the recommendation is only marginally worse then her most favoured one, she will never even know that it was not the most optimal (from the user's perspective) product that was recommended to her.

3 Towards prudent music recommender systems

Although it is difficult to come up with a single policy response that would reduce the negative consequences of RS bias, it is possible to identify a set of possible policy responses. Some of these are already in place (for example in the EU), or at an advanced stage of development.

A typical policy approach is transparency. This can take two complementary forms. First, transparency can be required around the RS design itself. To this end, the EU Platform To Business Regulation (Regulation (EU) 2019/1150) places a number of transparency requirements on intermediation platforms.[2] Similarly, the proposed EU Artificial Intelligence Act specifically requires transparency obligations for systems that interact with humans. Such transparency is potentially valuable in enabling users to understand how much reliance they can place on the recommendations they receive (Sinha & Swearingen 2002). It is arguable that such transparency is already required under existing consumer protection law. In a similar vein, the UK Competition and Markets Authority has accepted commitments from a number of hotel online booking sites to improve clarity around their rankings and issued principles for the sector.[3]

[1] `https://en.agcm.it/en/media/press-releases/2021/12/A528`.

[2] See Article 5(5).

[3] Competition and Markets Authority, Online hotel booking, GOV.UK (Oct. 27, 2017).

Transparency can be required around the actual outcomes of the RS. For example, where music companies have paid for higher rankings or better positioning, it could be argued that this constitutes advertising and thus that any such items should be clearly labelled as such. Such labelling should be sufficiently clear and prominent that consumers are readily able identify these items as paid advertising. Transparency requirements may also be limited to simply disclosing which objectives the RS' is maximising (e.g. whether revenue is part of it). This could help regulators better understand where incentives lie. For example, a platform is more likely to engage in self-preferencing behaviour if the RS is seeking to optimise the joint revenue of the platform, than if it is seeking to maximise only the utility of the users.

Moreover, given the discussion of biases above, it may be appropriate to require the RS to assess the levels of bias arising, and take proportionate steps to minimise it. For example, platforms could reduce the bias by improving RS design by applying hybrid models. Introducing serendipity into the objective function can also help recommend more products from the long-tail. This distributes revenues more evenly between suppliers. Most good RS will be carrying out this sort of ongoing assessment and adaptation process anyway, but others may need the motivation of a regulatory requirement. In any case, whether platforms in fact have the incentives to ameliorate their biases will presumably be influenced by the competition they face from other RSs on other platforms, and any benefits they may gain from the competition implications of the above biases.

Finally, it is important that public authorities are in a position to enforce and monitor whatever requirements are put in place for RS. This requires both ensuring relevant expertise within these public authorities and that they have access to private information about the RS design, including the objective function and any internal A/B (or other) testing that is carried out to assess outcomes.

References

Azaria, A., Hassidim, A., Kraus, S., Eshkol, A., Weintraub, O. & Netanely, I. (2013), Movie recommender system for profit maximization, *in* 'Proceedings of the 7th ACM conference on Recommender systems', pp. 121–128.

Bourreau, M. & Gaudin, G. (2018), 'Streaming platform and strategic recommendation bias'.

Brynjolfsson, E., Hu, Y. & Simester, D. (2011), 'Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales', *Management Science* **57**(8), 1373–1386.

Chaney, A. J., Stewart, B. M. & Engelhardt, B. E. (2018), How algorithmic confounding in recommendation systems increases homogeneity and decreases utility, *in* 'Proceedings of the 12th ACM Conference on Recommender Systems', pp. 224–232.

Ciampaglia, G. L., Nematzadeh, A., Menczer, F. & Flammini, A. (2018), 'How algorithmic popularity bias hinders or promotes quality', *Scientific reports* **8**(1), 1–7.

Holtz, D., Carterette, B., Chandar, P., Nazari, Z., Cramer, H. & Aral, S. (2020), The engagement-diversity connection: Evidence from a field experiment on spotify, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 75–76.

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

Schrage, M. (2020), *Recommendation Engines*, MIT Press.

Sinha, R. & Swearingen, K. (2002), The role of transparency in recommender systems, *in* 'CHI'02 extended abstracts on Human factors in computing systems', pp. 830–831.

Vall, A., Quadrana, M., Schedl, M. & Widmer, G. (2019), 'Order, context and popularity bias in next-song recommendations', *International Journal of Multimedia Information Retrieval* **8**(2), 101–113.

Waldfogel, J. (2017), 'How digitization has created a golden age of music, movies, books, and television', *Journal of economic perspectives* **31**(3), 195–214.

Zentner, A., Smith, M. & Kaya, C. (2013), 'How video rental patterns change as consumers move online', *Management Science* **59**(11), 2622–2634.

Zhang, L. (2018), 'Intellectual property strategy and the long tail: Evidence from the recorded music industry', *Management Science* **64**(1), 24–42.

Zhang, X., Ferreira, P., Godinho de Matos, M. & Belo, R. (2021), 'Welfare properties of profit maximizing recommender systems: Theory and results from a randomized experiment.', *MIS Quarterly* **45**(1).

Zhao, Z., Chen, J., Zhou, S., He, X., Cao, X., Zhang, F. & Wu, W. (2021), 'Popularity bias is not always evil: Disentangling benign and harmful bias for recommendation', *arXiv preprint arXiv:2109.07946* .