

d'Andria, Diego

Article — Published Version

Insider imitation with product differentiation

Journal of Evolutionary Economics

Provided in Cooperation with:

Springer Nature

Suggested Citation: d'Andria, Diego (2025) : Insider imitation with product differentiation, Journal of Evolutionary Economics, ISSN 1432-1386, Springer, Berlin, Heidelberg, Vol. 35, Iss. 1, pp. 95-122, <https://doi.org/10.1007/s00191-025-00886-w>

This Version is available at:

<https://hdl.handle.net/10419/323199>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Insider imitation with product differentiation

Diego d'Andria¹

Accepted: 7 January 2025 / Published online: 7 February 2025
© The Author(s) 2025

Abstract

Online marketplace providers such as Amazon may also offer own-brand products, in this way competing with third-party traders on their platform. Moreover they may exploit nonpublic data on third-party sales to identify opportunities for profit and design their own products, thus engaging in a form of partial imitation that is exclusive to them. Such an asymmetry in the ability to imitate has led to advocating for stricter regulation of online platforms in order to limit their ability to exploit nonpublic data. I propose a dynamic simulation model to study the effects of insider imitation by a platform provider under product differentiation. Using different parametrizations, I am able to show how such a market may evolve in time with or without the presence of insider imitation under different initial conditions. The model also allows to account for subsequent innovation via imitation (which I refer to as copy-and-be-copied dynamics), an aspect not considered by previous studies. These results add to the existing literature and cast further doubts on current regulatory approaches aiming at limiting data exploitation. Results suggest that the presence of insider imitation is unlikely to lead to smaller consumer welfare and larger market concentration and should not be banned on the grounds of standard antitrust concerns.

Keywords Amazon · Digital markets · Imitation · Marketplaces · Product innovation · Sherlocking

JEL Classification D40 · D82 · D83 · L40 · O3

1 Introduction

Marketplaces that are operated as digital platforms are serving an ever-enlarging share of trading companies and consumers worldwide. The largest, most famous and most influential of such platforms is Amazon. According to data reported in Faherty et al. (2017), yearly e-commerce retail sales in 2016 were 8.1% of total retail sales in the US and, out of US e-commerce retail sales, the domain Amazon.com alone managed

✉ Diego d'Andria
dandria@hs-sm.de

¹ Schmalkalden University of Applied Sciences, Schmalkalden, Germany

about 43% of them. More recent data suggest similar market shares are observed in recent years. Academics, policymakers and the general public have consequently been paying increasing attention to online marketplaces in general, and to Amazon in particular.

As members of the so-called Big Tech companies, some platform providers such as Amazon and Apple have been put under scrutiny by the antitrust authorities in the U.S. and in Europe. Recently the European Commission opened a new and more focused investigation related to the use of data by Amazon, in its dual role as platform provider and seller of own-brand products on the Amazon marketplace: “On 17 July 2019, the Commission opened a formal investigation to assess whether Amazon’s use of non-public data from independent retailers selling in its marketplace breached EU competition rules. On 10 November 2020, the Commission issued a Statement of Objections outlining its preliminary view that Amazon should not rely on independent sellers’ business data to calibrate its retail decisions, as this distorts fair competition on its platform and prevents effective competition” (European Commission 2022). In order to address these concerns, Amazon has committed “[...] to refrain from using nonpublic data relating to, or derived from, the activities of independent sellers on its marketplace, for its retail business that competes with those sellers.[...] Amazon commits not to use such data for the purposes of selling branded goods as well as its private label products.”

This novel antitrust approach to online marketplaces deserves specific study. The use of nonpublic data collected by a platform provider on their platform and their exploitation to perform insider imitation with own-brand products is the object of this paper. I argue that the solution agreed with Amazon to cope with the problem raised by the European Commission in its latest investigation, may be harmful for consumers and even lead to increased market concentration. I focus on analysing market outcomes in the form of total consumer welfare and market concentration. The analysis is driven by the desire to shed light on a very specific policy action, namely the mentioned investigation by the European Commission opened on 17 July 2019. While large platform providers such as Amazon may be harming market competition in many different ways (e.g., by exploiting network externalities or granting privileges to their own brand products via their internal search engine and interface), I concentrate my attention solely on their use of private data collected on their own platform and their exploitation for offering own-brand products.

Anticipating one of the main results stemming from the analysis, my simulations suggest that under most circumstances, insider imitation does not decrease consumer welfare. In some specific cases, insider imitation may even improve welfare and this would mean that preventing it (via regulation or antitrust action) might be welfare-deteriorating for consumers. The key mechanism here is the fact that, if one postulates that the platform provider may copy products on its platform, third-party sellers are able to do the same. Put in other words, a symmetrical condition must hold with respect to the inability to protect one’s own intellectual property for all participants to the market: if the platform provider at some point in time copies some products features, third sellers may copy them from the platform provider in subsequent times. The latter observation implies that although a platform provider may be able to exploit a larger information set to inform its product imitation choices, its appropriation of

the corresponding monopolistic rents will be at some point challenged by third-party introduction of new products. Such copy-and-be-copied behavior leads to a dynamic market evolution where the presence of insider imitation may, or may not, be welfare-deteriorating in the long run, based on whether the displacement effect of third-party products is stronger than the productivity-inducing effect due to faster product discovery in the next periods. A different way to look at insider imitation is as follows. The platform provider uses a larger information set compared to any third-party seller. As such, it gains an advantage in the race to discover the most profitable products. But, because products may be copied by others, the additional information generated by insider imitation remains and is used by all market participants, while the first mover advantage of the platform provider is eroded over time by imitators.

A relatively recent literature focuses on e-commerce platforms. Although some aspects of such platforms may be modelled as forms of two-sided markets (à la Rochet and Tirole (2006)), some specificities require dedicated analysis. An instance of the latter is the possibility that the platform provider also enters a market as seller of own-brand goods, thus competing with other traders on the platform. Etro (2021, 2022) and Hagiú et al. (2022) study the set of incentives faced by a platform provider and the conditions under which its entry, in the form of vertical integration, may be welfare-enhancing or deteriorating for consumers, considering that such providers face a trade-off between the prospect of profit from selling own-brand products and the loss of commissions paid by displaced competitors. Crawford et al. (2022) offer an empirical study by collecting data from the Amazon marketplace's Home and Kitchen department in Germany between 2016 and 2021. The results in Crawford et al. (2022) do not seem to provide support to the antitrust actions mentioned earlier. In the authors' words: "Our results provide much stronger evidence for internalizing platform externalities (to the benefits of consumers and third-party merchants alike) than expropriation of third-party seller sales". One of the reasons why the literature has gained interest in studying platforms is related to policy-making. In Gutierrez (2021) the author presents an empirically estimated model with the aim to better understand how specific policies restricting Amazon in a way or another, may affect consumers' welfare and competition on the platform, again obtaining mixed results. The question whether the use of own brand products to displace competitors may lead to stronger monopolistic power for the platform provider is highly debated (e.g., see Khan (2016) and Faherty et al. (2017), both arguing for stricter regulation of the platform providers), however both theoretical and empirical results from the economic literature are at best mixed and, so far, do not provide strong support for such policy actions.

The focus in this paper is closest to Madsen and Vellodi (2024). In Madsen and Vellodi (2024) the authors develop a theoretical model to study the possibility that a platform provider exploits data generated on the platform, with the aim to detect those sub-markets where entry has the largest profit potential (a practice that has been observed on multiple digital platforms and is sometimes referred to as *Sherlocking*: Choi et al. (2023)). In this way, the platform provider would engage in a form of competition that might reap innovation returns and thus reduce subsequent incentive for firms to innovate their products. The analysis tackles the question of whether exploitation of this specific use of data should be regulated. A limitation of Madsen and Vellodi (2024) is that their model assumes one-dimensional product innovations,

meaning that the platform provider imitates existing products by entirely copying them (“copycat” imitation), thus ruling out any possibility for partial imitation. I argue here that a platform provider is able not only to identify existing products that are more profitable, but also individual product characteristics that may be copied, mixed and matched to obtain new products that are not already traded on the marketplace. In this way, I model a form of creative imitation as defined in Kim and Nelson (2000) and Wang et al. (2023).

The model presented herein differs from Madsen and Vellodi (2024) in several ways. First, as stated, companies in my model engage in partial imitation rather than in exact imitations of existing products. Second, innovation is effortless while in Madsen and Vellodi (2024) it entails some investment in R&D. Third, Madsen and Vellodi (2024)’s methodological approach is an analytical non-computable model where an individual third-party seller innovates and is then potentially copied by a platform provider. The model presented herein is agent based, meaning that a computable model simulates the interaction between many third-party sellers, consumers and possibly a platform provider. Thus, my analysis complements the one in Madsen and Vellodi (2024) and provides an evolutionary perspective on the effects on welfare, market concentration and rate of innovation that insider (partial/creative) imitation is likely to generate. The dynamic nature of the model also allows to represent the possibility of an imitator who imitates the platform provider, thus also accounting for potentially important effects (in terms of “copy and be copied” dynamics as explained above) that entry of a platform provider with own products may have on subsequent innovation and that were not addressed in previous literature.

For illustration purposes, in order to provide the reader with a concrete example of how partial imitation may be shaped by nonpublic data collected by a platform provider, in the Appendix I describe the case of cloth hangers. When searching for cloth hangers on the Amazon.com website, its internal search engine returns several different products sold by different companies. The array of characteristics defining each product is large: the material the hanger is made of (wood, plastic, metal...), its geometry, the color and presence of ornamental design, the possibility to rotate the hook, presence of clips, additional hooks, anti-slip shapes or coatings, and more. These individual product characteristics can be mixed and matched together by any manufacturer, however the platform provider is the only agent gifted with the ability to data-mine a large number of transactions on their own platform, in order to discover which characteristics are statistically associated with better market outcomes. While it is likely that platform providers already use this form of data analysis (Crawford et al. 2022 report that almost 50% of own-brand products by Amazon in their sample were never offered before by other merchants, thus suggesting that partial innovation is indeed already an important part of Amazon’s own-product strategy), it is conceivable that in the near future application of more sophisticated deep learning algorithms will improve it even more. In particular, the use of artificial intelligence (AI) technologies may automatize both the recognition of distinct product characteristics (e.g., via image recognition and/or interpreting the text describing each product, a feature that modern generative AIs are already offering) and the subsequent data analysis (by applying standard econometric techniques or machine learning like Naive Bayesian methods). It is therefore important to gain proper insight into how insider trading performed

by platform providers via such data mining may shape competition and consumer welfare, also with the objective to provide antitrust authorities with a set of tools to analyse practical cases, and policymakers to design better market regulations.

To this end, I develop an agent-based model that represents product discovery in an evolutionary way. The model allows to take into account the dynamic nature of discovery and competition, together with the local nature of interactions between consumers and products (as consumers usually make their purchase decisions based on a limited number of observed products). Agent-based modelling techniques have some distinctive advantages and drawbacks when compared to non-computational modelling (Tesfatsion 2003; Tesfatsion and Judd 2006), one of their strengths being the possibility to represent heterogeneous agents interacting locally in the presence of path dependencies. My approach draws inspiration from evolutionary theories of innovation (Nelson and Winter 2009; Cantner and Hanusch 2002). As concrete examples of applying such theoretical framework to the analysis of a specific market, see Cantner (2007) and Cantner et al. (2019)) and from studies that model product spaces to analyse competition among firms with differentiated products and product innovation (Dawid 2006; Kiesling et al. 2012). The present model focuses on innovation diffusion in a product space, in this being close to Ma and Nakamori (2005). Contrary to other studies (e.g., Yildizoglu (2002) and d'Andria and Savin (2018)) that study market competition in an evolutionary fashion, I do not explicitly model R&D investment as I simply assume a common production technology available to all competitors. Product innovation in the model is the sole outcome of a recombination of product characteristics and exploration of different bundles of these, the success of which may in time lead to gaining market shares and to the diffusion of a given bundle of characteristics via costless imitation. Thus, any gained advantage by a firm is solely driven by a better fit of a product to consumers' preferences and is never due to productivity or cost advantages.

In the model each period features a random sample of consumers with heterogeneous preferences, who are able to observe a limited number of products available on an online platform. Consumers are characterized by a utility function, made of a quasilinear function in which a Cobb-Douglas aggregation of two CES functions is nested. This utility function is very flexible as it is characterized by a set of parameters which allow different consumers to attach not only different importance to each individual product characteristic, but also different degrees of complementarity or substitutability to pairs of characteristics. Firms may engage in product innovation by either experimenting with some entirely random innovation in product characteristics, or by imitating some characteristics from the best-selling products. The model then allows, optionally, to assume the existence of one additional firm, which in the following I label as the Platform firm, that may also engage in partial imitation where such imitation process is informed by data collected on all transactions in the last simulated period. By using different sets of parameters I am able to study the conditions under which the presence of an Platform firm is eventually associated with lower welfare, higher market concentration, smaller variety of traded products, and stronger displacement of competitors' profit to the benefit of the Platform firm.

The model is highly stylized, in the sense that it disregards firm-specific characteristics. Any seller who participates in the simulated market has the same *a priori* chances to discover a product and sells it at the same price as any other competitor would. When a firm is “large” in the model, it only means that such firm offers a wider range of product varieties in this market, and/or that it produces a larger profit or turnover. A firm, however, in this model can never cross-subsidize products or displace competitors via price or product bundling strategies, nor can it enjoy any productivity advantage. In this way, I eliminate a number of potential confounding factors and solely focus on the combined effects that product innovation via partial imitation, limited information and the possible existence of an insider imitator produce.

The following sections are as follows. Section 2 introduces the reader to the main assumptions of the model and illustrates how the simulations work. Different scenarios are detailed to capture different possible states of the market. Section 3 discusses the main results by comparing average outcomes across a large number of simulations. Section 4 concludes and draws relevant implications for future research, policymakers and antitrust practice.

2 The model

The model represents a single market. All products offered are represented by a vector \underline{v} of fixed length n of characteristics, each characteristic ranging between zero and one. To provide an intuitive mapping of this product space to an actual product, consider the example of a chair. The market for chairs is made of different varieties, where each product can vary based on materials, design and a number of features (armrests, amortizers, wheels, etc.). A characteristic in this product space might be, for example, “durability” ranging from an extremely fragile product (corresponding to a value for said characteristic of zero) to an extremely robust one thanks to the use of sturdier materials (corresponding to a value for said characteristic of one).

In summary, the model is simulated, first, by generating a population of consumers and a population of firms. Then, in each simulation period, some consumers observe some products and decide whether to buy them. At the end of a period, each firm may choose to perform product innovation and add a new product to the range offered, and may discard one unprofitable product. Past events therefore determine if and how product innovation occurred, the range of products available in the market and, based on the random matching between few consumers and few products, the information available to firms in order to make their choices about new products to offer in the next period.

2.1 Population of agents and initial setup

There is a fixed number of consumers in the model. Each consumer is characterized by a quasilinear utility function of the form:

$$U(\underline{v}) = (A^\beta B^{1-\beta})^{0.5} + W - p(\underline{v}) \quad (1)$$

where W is income (assumed to be the same for all consumers), $p(\underline{v})$ is the price of a product having characteristics \underline{v} (I use the underlined font to indicate a vector: $\underline{v} = \{v_1, \dots, v_n\}$), and A and B are two nested CES functions:

$$A = (\sum_{i=1}^m c_i v_i^r)^{1/r} \quad (2)$$

$$B = (\sum_{j=m+1}^n d_j v_j^s)^{1/s} \quad (3)$$

with $0 < r < 1$, $s < 0$, $\sum c_i = 1$ and $\sum d_i = 1$.

Utility is expressed using a novel combination of well-known functional forms: a parent quasilinear function and, nested into it, a Cobb-Douglas function. Inside the Cobb-Douglas function, furthermore, are nested two Constant Elasticity of Substitution (CES) functions named A and B and detailed in Eqs. 2 and 3. Inside Eq. 1 the addendum $W - p(\underline{v})$ represents spending money on any other good or service available in the economy, other than the one traded on the simulated online market. It may be considered as spending on a Hicksian good with price normalized to one. The use of a quasilinear function is convenient as changes in utility can be directly translated into monetary values, which simplifies welfare comparisons.

The second-level functional form in Eq. 1, as stated, is a Cobb-Douglas function in which two CES functions are nested. The first CES is named A and accepts as arguments individual product characteristics and has an elasticity parameter, such that these characteristics act as substitutes, meaning that increasing one of its arguments may not increase utility (nor decrease it). The other CES function, named B , is parametrized such that its arguments act as complements, therefore increasing the value of one may increase utility in a faster way (due to the multiplicative effect of its complementary characteristics). All arguments within A and B are weighted (these weights are indicated as c_i and d_j in Eqs. 2 and 3), and these weights are also randomly assigned to each consumer. This parametrization of the two nested CES functions implies that the characteristics represented within A act as substitutes, while those in B act as complements. By randomly assigning characteristics to either A or B and randomly determining the weights c_i and d_i , the model accommodates a large variety of heterogeneous preferences with regard to a product's features. The two elasticity parameters, r and s , are also randomly assigned to each consumer from a uniform distribution, strictly bounded respectively between zero and 1, and between minus infinitum and zero. The parameter β in the Cobb-Douglas function merely acts as a scaling factor: values of β closer to one imply that consumers are more ready to buy an observed product, in comparison to the choice not to buy and spend money on the Hicksian good instead. In the simulations, this parameter is always set to 0.5. Overall, this choice of a functional form obtains a utility function that is continuous, monotonically non-decreasing in each of its arguments and guarantees that utility is maximized for some vector of characteristics that is not a corner solution (not one with all characteristics equal to their maximum possible value, 1).

There is a fixed number of firms. Each firm offers a number of products in the first simulated period, and this number is randomly drawn from a Pareto distribution with probability density function:

$$f(x) = \frac{\alpha m^\alpha}{x^{\alpha+1}} \quad (4)$$

The parameter α determines the shape of the distribution, particularly the thickness of its tail. An $\alpha = 2$, as assumed in the simulations, obtains a very close approximation of a 20%-80% distribution meaning that the initial number of products per firm is highly skewed, in this following the widespread observation that in most markets, suppliers tend to distribute according to an 80-20 rule (which means that about 80% of the occurrences are in the lowest 20% of the range). The parameter m represents the smallest possible value that the random variable can take, which in the simulations is set to 3, meaning that firms will start a simulation by offering at least three different products. Again, each product is defined according to a fixed-length vector of characteristics \underline{v} as explained before, in this matching the number of characteristics entering the utility function in Eq. 1. In the simulations, all vectors \underline{v} are made of ten characteristics. Each of them, labelled v_k (where therefore v_k indicates the k -th of the characteristics in vector $\underline{v} = \{v_1, \dots, v_n\}$) may range from zero to 1. When products are initially generated, the numerical value assigned to each v_k is individually drawn from a uniform distribution between 0 and 1. Profit for a firm selling L different product varieties is:

$$\pi = \sum_{l=1}^L (p_l - \sum_k c(v_k))_l \quad (5)$$

Thus, $\sum_k c(v_k)$ is simply the sum of the individual costs associated to each characteristic in \underline{v} . Here the production cost of an individual product is determined by its \underline{v} characteristics. For each single characteristic v_k in \underline{v} , I assume a quadratic cost function $c(v_k) = \frac{v_k^2}{2}$. All firms have access to the same production technology. Sale price is obtained by applying a fixed mark-up ϕ on costs, also assumed to be equal across all firms, such that the sale price is $p_l = (1 + \phi) \sum_k c(v_k)$. Thus, a product offering better quality in some areas will cost more to produce than a lower-quality one, but it will also sell for a proportionally higher price.

At the time when the simulation is initialized, a population of consumers is randomly drawn where the values for r , s , the weights c_i and d_i are obtained from the respective uniform distributions. A population of firms is randomly generated by using the distribution under Eq. 4 to determine the number of products offered in the first period. Firms are also randomly assigned a type at the beginning, which can be either “explorer” or “imitator”. The latter attribute only affects how a firm may introduce new products (while for all other aspects, all firms behave in the same way). Optionally the model allows for a third type of innovation, which is exclusive to a Platform firm (more on this later). In this way, each replication of the simulated scenario is performed on a different initial set of consumers, firms and offered products. While consumer preferences will remain fixed throughout all simulated periods, firms will be able to change their products via innovation (see below).

2.2 Trading

The model is simulated for subsequent periods of time. In each period, a random sample of consumers is allowed to trade. Each of these consumers is allowed to see a limited number of products available on the market, which are selected among the

available ones by applying a ranked search algorithm (more on the latter in Section 2.5). This limitation on the ability of consumers to compare available products is meant to represent the fact that, on average, consumers on online marketplaces are found to spend only few minutes per product search. In the baseline simulations 250 consumers out of the 1000 initially generated are allowed to trade, and each of them can see at most 30 products offered on the market. For each observed products, the consumer-specific utility function under Eq. 1 determines a ranking of products and whether the most-preferred one will be bought. Note that the quasilinear form given to utility and the chosen parametrization for the simulations imply that not all consumers will buy a product, as they will only enter a purchase if the product bearing the largest utility for them among those observed in the period, offers them a utility gain compared to just keeping their income (which would ideally be spent on some Hicksian composite good that is not explicitly modelled in this setting). Also note that the fact that consumers only evaluate a finite number of products in each period is a crucial element differentiating it from more traditional microeconomic models where consumers observe all available products at once. Firms are assumed only to bear variable costs (refer again to Eq. 5) and are able to satisfy any number of purchases for any of the L products they offer. At the end of a simulation period, each firm will have generated a profit according to Eq. 5.

2.3 Product innovation

At the end of each simulated period, after all trades have occurred, each firm may innovate. Innovation is stochastic and each firm has the same chances (30% in the Baseline) to introduce a new product, regardless of their type and performance. By innovation here I refer to the introduction of a new product, that was previously not offered on the market. The way in which a firm innovates differs based on their type, assigned at the beginning of a replication (in my parametrizations, 50% of the firms belong to one type and 50% to the other). The two types of firms are labelled explorers and imitators for easier reference.

“Explorers” innovate by adding a randomly mutated product to the array of products they offered in the last period. The choice for such new product is totally disconnected from the observed behavior of consumers and competing products. As such, explorers are meant to represent firms who try to introduce radically new products into the market. When an explorer innovates, the algorithm proceeds as follows:

1. A random number of characteristics to innovate is determined (between 1 and 4, out of the 10 characteristics possessed by each product). Each defines a position in the product vector \underline{v} (thus, from position 1 to position 10).
2. One of the existing products of this explorer firm is randomly selected.
3. The 1-4 chosen characteristics from step (1), out of the ten contained in vector \underline{v} , are substituted with a random draw from the uniform distribution between zero and 1.
4. The resulting new product is added to the array of products offered by this explorer firm.

Thus, explorers generate new products without using any information from past trades or present market conditions: They proceed following a trial-and-error strategy, starting from their own array of products and then randomly exploring new product directions in the space of all possible product characteristics.

“Imitators”, on the contrary, try to exploit information obtained from the market. Imitators observe some of the best-performing products and copy ideas from them. Each imitator randomly chooses one of the products that performed best in the last period (best-performing here means that these products produced the largest total profit for the seller), out of the five best-performing products market-wise. Imitators randomly select a number of characteristics from these products and substitute them into one of their own products, thus giving birth to a mutated variety that is a hybrid between an own product and the market best-performing ones. Consequently, “Imitators” engage in a form of partial imitation where they are able to only observe some of the best-performing products from the last period and, out of them, try to guess which specific product characteristics are the ones that mostly drive such commercial success. Their newly introduced products are therefore crossovers between existing products. More in detail, the algorithm works as follows:

1. A random number of characteristics to imitate is determined (between 1 and 4, out of the 10 characteristics possessed by each product). Let us call this number N . Thus, only few aspects of a successful product are injected into another product of the imitating firm. Imitation therefor never entails 100% copying an existing product.
2. One of the five best-performing products in the market from last period is randomly chosen, from which one of its ten characteristics is randomly selected. This is repeated up to three times more thus obtaining N characteristics to copy from (each in a different position in the vector \underline{v} defining a product, and each from any of the five best-performing products).
3. An existing product from the ones offered in the last period by this imitating firm is randomly selected.
4. The N copied characteristics are substituted into the latter product (from step (3)).
5. This new product is added to the array of offered products of the imitating firm starting from the next period.

As stated, both explorers and imitators are allowed to select a random number of characteristics to mutate, up to a predefined maximum number. It is to note that both explorers and imitators are unable to understand which characteristics are driving the success of a product. Therefore, imitators pick some characteristics from best-performing products without having any understanding whether these are indeed the ingredient determining the success of the latter. One interesting feature of this modelling strategy is that imitators are able to imitate from imitators. This dynamic mechanism is, in this author's view, of great importance to shed light on the presence of an insider imitator that may exploit private data for its own imitative process. Third-party imitators may imitate the Platform firm, among others, thus the presence of this special agent affects not only its own profitability, but also all other imitators' subsequent innovations as they over time observe best-performing products and try to copy from them.

Firms not only may add new products, they may also drop old ones. Firms whose products were particularly unsuccessful will gradually reduce the range of products they offer: each period, each firm has a 50% chance to drop one product that earned no profit in the last period.

2.4 Insider imitation

The model allows for an optional type of firm, which throughout the rest of the paper is labelled as the “Platform firm”. This special firm innovates in a unique way: it adds new products as imitators do, with the difference that the mutated characteristics are driven by an Ordinary Least Squares (OLS) estimation run on the entire market, rather than randomly selecting characteristics from the best-performing products.

In each period, a firm of type “Platform” is given the same probability to innovate as explorers and imitators (as stated, this is 30% in the Baseline). If it innovates in a period, an algorithm is executed which performs the following steps:

1. Observe all trades that happened in the last period; sum up profit generated by each individual product.
2. Perform an OLS estimation on the model: $\pi \sim v_1 + \dots + v_n + v_1^2 + \dots + v_n^2$. Here π stands for total profit generated by a product in the last period (obtained in previous step (1)), while v_1 to v_n are the characteristics of that product from vector \underline{v} (note that the choice of adding the characteristics also as quadratic terms in the OLS model is meant to capture internal solutions to the profit optimization problem). An optimizer is then used to find the theoretical combination of characteristics \underline{v}^* that would maximize π , based on the estimated OLS coefficients.
3. Randomly select a product from the five best-performing products on the market; randomly select N characteristics to mutate (1 to 4 characteristics, as explorers and imitators do).
4. Modify the N selected characteristics so that they are increased (or decreased) by a random addendum (randomly chosen from a uniform distribution between the imitated characteristic's current value and the theoretical one predicted by the OLS estimation under previous step (2)). This step is performed in order to bring characteristics closer to the theoretical value derived from the OLS estimated coefficients, which would maximize expected profit.
5. Finally, add this new product to the ones offered by the Platform firm in the next period.

The special advantage given to the Platform firm in the model is that:

1. Compared to explorers, the Platform firm exploits market information. Explorers generate new products in a totally random and uninformed way.
2. Compared to imitators, the Platform firm is able to exploit more information. Imitators only know which products performed best in the last period. The Platform firm also obtains information about correlations (via the OLS approach) between individual values of characteristics across all traded products and their produced profit.

By means of applying OLS to all transactions in a given period, this special firm engages in a kind of directed search in the space of product characteristics which has better chances to produce a more profitable product, compared to “imitator” firms. At the same time, the Platform firm does not engage in the more radical and non-directed experimentation that “explorer” type firms perform.

Even though the Platform firm is gifted with enhanced insight, its product imitation is imperfect: information comes from observing recent transactions which involve only a fraction of existing consumers, thus the Platform firm cannot truly engage in a full-fledged profit optimization as aggregate demand changes at each iteration. This is the reason why imitators may improve over the Platform firm's products by discovering some bundle of characteristics that is even more profitable in a given period of time. By the same reasoning, explorers may too discover more profitable bundles, especially in the early periods of a simulation when the market has not yet converged close to the most attractive bundles for consumers. This noisiness in obtained information adds up to the fact that differentiated products and heterogeneous consumer preferences may be such that multiple profit-maximizing maxima exist and await to be discovered. These maxima are however always function of current market conditions and depend, too, on the specific range of products offered by all competitors in a given period. Put in other words, the profit-maximizing product that could be offered in the market is state-dependent, and expectations about it are also path-dependent.

It is worth stressing that the Platform firm is able to appropriate rents from its innovation efforts exactly like all other firms in the market. This means that, based on the speed at which the main features of its successful new product are subsequently imitated by other firms (where such speed is determined stochastically in the model), such rents can last for few or for several simulated periods. This is in line with the available evidence (refer again to Crawford et al. (2022)), according to which many own-brand Amazon products are only offered for a limited time while others persist for several years.

2.5 Ranked search results obtained by consumers

In the base model, the results shown to a consumer are ranked, such that a product that has been more successful in the last period has higher chances to be observed by consumers. This is referred to as ranked search algorithm. The latter is meant to represent consumers who mostly search for a product via a marketplace platform's internal search engine, such that products with more and/or better evaluations are reported more often than products with worse and/or fewer ratings. This form of ranked search is akin to having a rating system in place, which as observed already (e.g., in Vellodi (2018)), may act as an additional barrier to entry that may also affect the success of a new product.

In ranked search scenarios, each consumer always sees a fixed number of products (30 in the Baseline). Half of these products are randomly selected from all available products, while the other half is obtained by a weighted random selection, with weights made equal to the share of profit generated by an individual product over total profit

in the market. In this way, consumers are more exposed to successful products while still being able to evaluate a few, less popular products.

It is to note that consumers may land on a product page in a given marketplace (e.g., the Amazon marketplace) also from the outside of it. A typical example is when a consumer searches for a product on a search engine (e.g., the Google search engine) and clicks on a returned suggestion which leads to a product on Amazon. Or, alternatively, a consumer might use a Web service listing products according to a given criterion (e.g.: the cheapest products across several marketplaces and direct e-commerce stores; the products rated highest by a committee, users' group or certification agency). Thus, modelling a scenario with ranked search is not always the one bearing the largest external validity as different marketplaces, and different categories of products for the same marketplace, may experience heterogeneous modalities of access. For the specific case of Amazon, available evidence suggests that currently a very large majority of consumers search for products using the internal search engine of the Amazon marketplace, which is the reason why ranked search is assumed in the base model. An optional model specification is simulated for a smaller number of periods (only 50) to offer a term of comparison and for added robustness of the results. In this optional model specification, each consumer is shown a fixed number of products (30 as in the Baseline), chosen completely randomly instead of being assigned probability weights. The results of these smaller-scale experiments are reported in the Appendix, Figs. 6 and 7.

2.6 Simulations

The simulation model is written and executed using Python 3.11.4 (packaged by Anaconda). I ran a large number of simulated scenarios: For each set of parameters defining a scenario, I ran 100 repetitions and took average values of the indicators of interest for each period.

The number of periods the simulation runs through is kept the same across all scenarios and was selected as follows. I rely on the following convergence criterion: stop a simulation run when the sum of welfare in the last 20 periods, X , and the sum of welfare in the periods from -40 to -20 (where period “zero” is the last period being simulated), Y are such that $(X - Y)/Y < t$. This tolerance threshold t has been set to 0.01% (one part on 10,000), meaning that a simulation stops when welfare only grows by extremely small increments, compared to the previous 20 periods. I ran the baseline scenario 100 times (50 with, and 50 without, the Platform firm) letting the simulation run until the convergence criterion was met, and obtained that the simulation meets convergence, on average, on period No. 105, with a standard deviation of about 23.3 (corrected using the Bessel's correction). Assuming a Normal distribution, about 95% of the simulations will therefore stop in a period smaller or equal to 152. I set therefore the parameter for the number of periods to 155 in order to obtain a fixed number of periods where a very large majority (more than 95%) of the simulations will have converged.

I refer to the Baseline scenario as one with the parameters listed in Table 1. Most scenarios are then derived by varying only one or two parameters while keeping all

Table 1 List of parameters with Baseline values

| Parameter name | Value | Description |
|------------------------------|-------|--|
| no_consumers | 1,000 | Number of consumers |
| no_firms | 50 | Number of firms |
| no_characteristics | 10 | Number of product characteristics |
| periods | 155 | Number of simulated periods |
| replications | 100 | Number of repetitions for each scenario |
| consumers_no_trading | 250 | Number of consumers trading in one period |
| copyfrom_bestn | 5 | Top n products imitated from |
| innovation_prob | 0.30 | Probability that a firm innovates |
| mark_up | 0.10 | Price mark-up |
| max_innovated_chars | 4 | Max. number of innovated characteristics |
| min_no_characteristics_inbin | 2 | Min. number of characteristics in A or B |
| min_no_prod_amazon | 3 | Min. number of products for Amazon |
| min_no_products | 3 | Min. number of products for each firm |
| pareto_coeff | 2 | Coefficient in the Pareto distribution |
| prob_drop_zeroprofit | 0.50 | Probability to drop one unprofitable product |
| products_observed | 30 | Number of products observed by a consumer |
| sh_explorers | 0.50 | Percentage of firms of type Explorer |
| size_amazon_init | 5 | Initial number of products offered by Amazon |
| W | 1 | Disposable income for each consumer |

others fixed.¹ The most important element in the context of the following analysis is to derive scenarios with the same set of parameters but differing only because a first set has no Platform firm, while the second set includes it. In this way, I may shed light on a number of effects that the presence of an insider imitator such as Amazon may have on this market.

The Baseline parametrization was obtained from the following rationales. The probability to innovate was made equal to 30% to match figures published by Eurostat (items *inn_cis12_bas* and *inn_cis12_spec*, for the years 2018-2020 and for all firms with at least 10 employees) which reported that about 30% of survey respondents to the Community Innovation Survey introduced at least one product innovation in the period. Income W was set equal to 1 in order to have, on average, about two thirds of the consumers who participate to trading buy a product in the very early periods of

¹ All of the parameters listed in Table 1 have been subject to a sensitivity analysis, the results of which are not reported for space reasons. Some tests involved simulations with fewer replications but more periods (up to 250) in order to ascertain that the results obtained for 155 periods were also representative of longer time horizons. Other tests changed a single parameter, by increasing it by +50% or decreasing it by -50% in comparison to baseline values, or changing it by another significant amount (e.g.: doubling or halving the number of consumers; doubling or halving the number of trading consumers per period; setting the Pareto coefficient to 3 or 1). The exponent used in the Cobb-Douglas part of the utility function, set at 0.5 in all the reported simulations, was also tested with values 0.25 and 0.75. The main results reported hereafter, which are that the presence of a Platform firm does not decrease consumer welfare nor increase market concentration, were always confirmed in all said tests.

the simulation and about 90% buyers in the very last periods, given the other Baseline parameters in Table 1. Firms whose products were particularly unsuccessful will gradually reduce the range of products they offer (each period, each firm has a 50% chance to drop one product that earned no profit in the last period), however one must stress that the model assumes a contestable market where new products can be continually introduced. The latter happens in two ways: by means of innovation, and adding randomly generated products to a firm when this falls below 3 offered products. Because each product's production and evaluation are independent from other products from the same firm (i.e., no cross-subsidization is allowed in the model), the latter may also be interpreted as entry of new competitors in the market, although no formal entry and exit are modelled. Innovating firms are allowed to imitate and mutate up to four characteristics (of the total 10 which define a product). This limitation is meant to reduce chances that the model gets stuck in a local optimum too early. By having only few (250), randomly chosen consumers participating to trading in each period and each allowed to observe only few (30) products to choose from, the model displays small-world properties as interactions happen locally. For each consumer, any two product characteristics can either act as complements or as substitutes within their utility function, with different intensities. This implies that in each period aggregate demand is different, randomly determined by the specific consumers who participate to trading. A parameter for the Pareto distribution of 2 determines a slightly different distribution of firms than the 80%-20% ratio (again, this is the distribution of the initial number of offered products by each firm), which is meant to reduce chances to start many simulations with few, extremely large firms which would reduce performed innovation to unintended levels (because each firm has a chance to innovate one product in one period, starting with few firms flooding the market with many products would constrain total innovation performed; this aspect will be addressed by specific scenarios where I shock the number of initial firms). The parameter labelled *min_noproducts* enters the Pareto distribution in Eq. 4 as the scaling factor *m*, which makes each firm initially offer at least a number of products equal to *min_noproducts*.

Most of the different simulated scenarios are obtained by changing one key parameter only. I simulated a Baseline with 50 firms, then two variations with 100 and 20 firms respectively. Having more firms implies more available products and also more innovation performed in each period. I ran scenarios where I increase or reduce the number of products that can be evaluated by a consumer in each period (respectively 15 and 60 products, from the 30 visible products in the Baseline) to represent different conditions with regard to the ability to compare more products at once and the willingness of the average consumer to commit more time to searching. The probability that a firm innovates its products is also tested, changing it from 30% in the Baseline to 15% or 60%. Note that the latter scenarios behave slightly differently from the ones where the number of firms is modified, because the number of initial products and the average number of products in the market is not directly affected by the probability to innovate (although over time, more innovation in the model implies that more products will be made available).

There are also scenarios that are not obtained by modifying one individual parameter but rather modifying multiple ones simultaneously or by adding special features to the model. All simulations are obtained first without, and then with, a Platform firm. The

ones with the Platform firm, as explained in Section 2.4, have one additional firm that performs a unique kind of innovation via imitation. The Platform firm always starts a simulation with 5 products, which is a value between the mean and median number of initial products assigned by the Pareto distribution in (4) to the other firms. One special scenario allows the Platform firm to invest in a very large number of products when initially entering the market (30 products, up from the 5 allowed in the Baseline) and also I assume that the Platform firm always keeps a larger number (at least 10, up from 3 in the Baseline) of own products throughout the simulated periods. The latter is what is referred to in the following as a “Huge Platform” scenario. The Huge Platform scenario is designed to represent a hypothetical “flood the market” strategy meant to displace competing products with a large range of own-brand products from the Platform firm. Note that although the Platform firm may have a dominant position in the market for retail trade as a whole, the latter does not automatically imply that it has a dominant position in all of the sub-markets served by its platform (the opposite is more often observed, as illustrated for example by the empirical investigation in Crawford et al. (2022)). Therefore, the “Huge Platform” scenario constitutes a special case and is presented hereafter only to provide insight about possible interactions of such a dominant position with the kind of mechanism that is the focus of this study.

3 Simulation results

In this section I illustrate some notable results stemming from a large number of simulations. Most of the results will be reported in graphical form to allow for a more intuitive interpretation. Because my interest lies in understanding the effects of a firm performing insider imitation, I mostly compare simulated scenarios that are identical, save for allowing or not the existence of a Platform firm. My simulated setting allows me to observe all endogenous variables, both at the market and individual level, in each simulated period. The following results provide a summary and a bird’s eye perspective.

As a preliminary note, it must be stressed that even though the Platform firm is given, information-wise, a superior algorithm to improve its products, it does not necessarily imply that it will end up being the most profitable firm. For illustration purposes, Fig. 1 compares side by side two examples taken from individual replications of the Baseline scenario, where the y-axis is for total profit generated by each of the most profitable five firms plus the Platform firm. In the first example (graph on the left), the Platform firm does not enter the top-five (except for short time windows between periods 60 and 80, and around period 150) and lags behind significantly for most of the simulated periods. In the second example (graph on the right) the Platform firm rapidly gains market shares, up to the point when it becomes the leading firm (approximately between periods No. 10 to 50), before losing ground against competitors in the periods after 50. In a limited time window between about periods 10 and 45, the Platform firm is leader with a visible distance from the second-best competitor, but does not manage to maintain durable leadership until the end of the simulation. A similar behavior to the graphs in Fig. 1 is observed in all scenarios, as innovation and path-dependencies make the market always contestable.

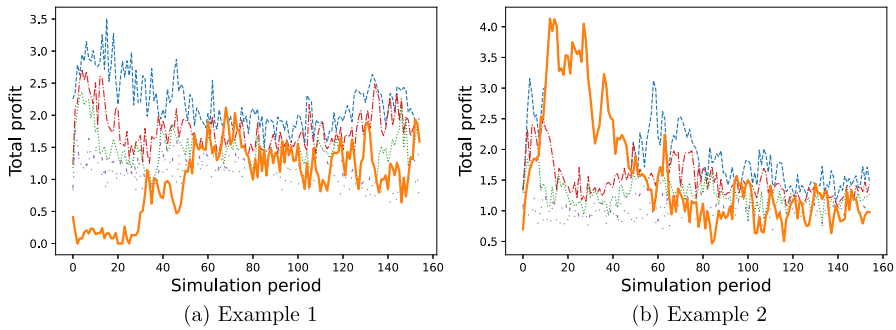


Fig. 1 Examples of two outcomes under Baseline scenario with the Platform firm. The y-axis reports total profit - Orange continuous lines are for the Platform firm, while the other five dotted or dashed lines are for the five best-performing firms

Welfare generated in the market is calculated by summing up, in each period, the utility obtained by each consumer (either by purchasing a product or not purchasing anything) plus profit earned by each firm. Figure 2 compares average welfare by plotting it for each scenario and, for each scenario, with or without the presence of the Platform firm. In each plot shown in Fig. 2 the blue dashed line represents mean values across 100 scenarios without the Platform firm, and the orange continuous line represents the same indicator across 100 scenarios with the Platform firm. Results point to an important conclusion: the presence of a Platform firm does not, per se, imply a loss in welfare and, under some specific parametrizations, it may even improve it. In the simulations, the only case where the presence of a Platform firm reduces welfare is the scenario with reduced probability to innovate (graph (f) in Fig. 2). Simulation results suggest that the number of firms is not a major factor affecting welfare, as the Platform firm is always found to improve welfare (albeit with different intensity) with 20, 50 and 100 firms. In the case of 20 firms, welfare with a Platform firm even increases in about the first half of the simulated periods. When consumers are strongly constrained in their ability to compare products or when firms (including the Platform firm) innovate at a slow pace, the welfare-enhancing effect of the Platform firm tends to disappear. In these simulations because of the assumption of a ranked search algorithm showing products to consumers, the market converges quickly to few high-value products and the positive welfare impact of the Platform firm disappears. In a set of alternative scenarios where consumers are instead assumed to observe products randomly (without ranked search features), the presence of a Platform firm more often is able to increase welfare (refer to Figs. 6 and 7 in the Appendix).

Also, as an aside note that because of the way this welfare indicator is built, a low product variety may mean more products are available which are more similar to each other, rather than fewer but more diversified products. Figure 5 (in the Appendix) reports indicators for the variety of products traded on the market, which is measured as the average standard deviation across all product characteristics and across available products that are traded at least once in the period. As the most profitable bundles of characteristics are discovered, firms tend to focus on them via imitation and hence, large deviations from such bundles become less common. In most tested scenarios the

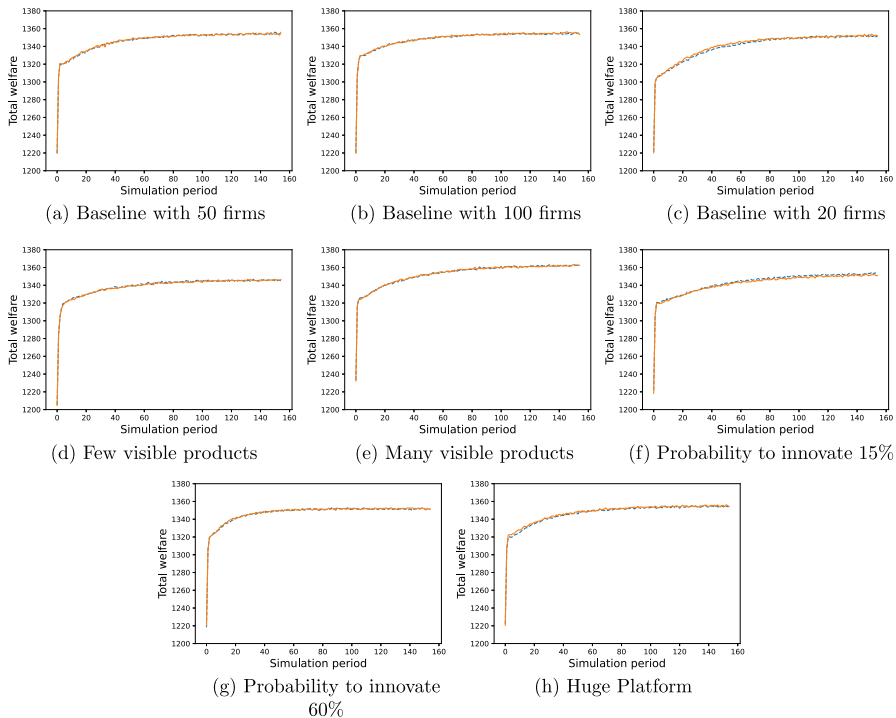


Fig. 2 Total welfare (average across simulated scenarios) - With a Platform firm (orange continuous lines) and without it (dashed blue lines)

presence of a Platform firm moderately reduces product variety. It is to note that I do not consider the case of consumers who derive utility from being able to choose from a larger range of different products. Therefore, different degrees of variety do not have any direct effect on welfare.

Figure 3 reports the Herfindal-Hirschman Concentration Index (HHI), which measures how concentrated the market is based on realized profit. Larger values of the HHI (closer to 1) indicate more concentrated markets. In light of the recent actions by antitrust authorities in Europe and the United States, this indicator is useful to shed light on the expected impact of insider imitation by platform providers on market concentration. An interesting and perhaps counter-intuitive result from the simulation analysis is that the presence of a Platform firm may both cause an increase or a reduction in market concentration. Moreover, in most simulated scenarios we observe a decrease in market concentration with a Platform firm, the only meaningful exception being scenarios with many visible products (see graph (e) in Fig. 3), which makes the Platform firm much better at innovating (thanks to its special imitation algorithm) which, in turn, earns the Platform firm a larger share of total market profit. Also, scenarios featuring a smaller probability to innovate or a Huge Platform produce mixed results with respect to the HHI. These results also imply that the impacts of the Plat-

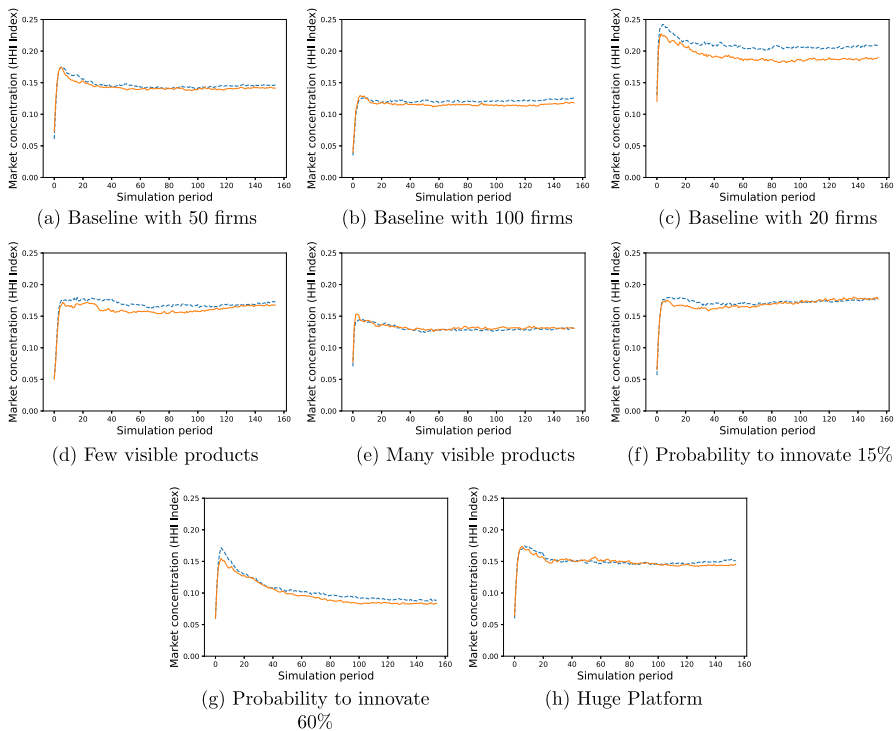


Fig. 3 Market concentration - Herfindahl-Hirschman Index (HHI, average across simulated scenarios) - With a Platform firm (orange continuous lines) and without it (dashed blue lines)

form firm on welfare and on market concentration do not necessary go in the same direction, as in some scenarios we observe improved welfare coupled with higher market concentration with a Platform firm, while in others we see improved welfare coupled with lower market concentration with a Platform firm.

The role of ranked search deserves a specific discussion. With ranked search, as assumed in all of the base model scenarios whose results are reported in Figs. 2 and 3, the Platform firm has significantly higher chances to reap larger profit in each period compared to scenarios without ranked search (as in Figs. 6 and 7 in the Appendix). However, profit earned by the Platform firm is still much smaller, on average, than what the first-earner firm reaps on average. Ranked search benefits all major competitors, not just the Platform firm. Again, as stated already, with ranked search the presence of a Platform firm is rarely found to increase market concentration, on the contrary it may sometimes reduce it.

Figure 4 focuses on the comparison of four specific scenarios: Baseline, scenario with more visible products, scenario with smaller chance to innovate, and the Huge Platform scenario. The latter three scenario are selected here because they produced somewhat different results than the generality of other results, either in the form of reduced welfare or increased HHI when a Platform firm is present. The Baseline

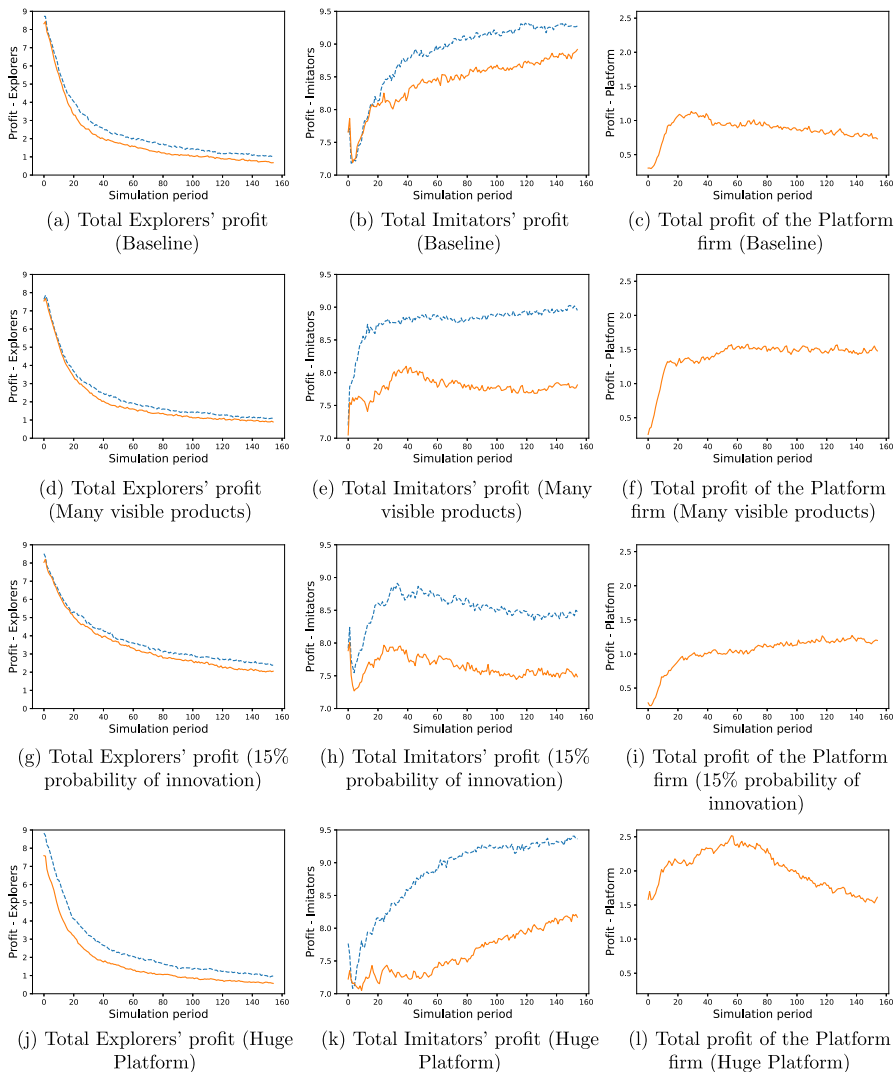


Fig. 4 Total profit earned on average by explorers, imitators and the Platform firm (Baseline vs selected scenarios) - With a Platform firm (orange continuous lines) and without it (dashed blue lines)

scenario behaves similarly to all other scenarios which were defined by only changing one single parameter and is, therefore, representative of a large set of possible parametrizations. The scenario with 15% probability of innovation is the only scenario that, in the simulations, obtained lower welfare in the presence of a Platform firm, thus it is relevant to further study it to better understand the possibility of a welfare-deteriorating effect of the Platform firm's insider imitation. In Fig. 4 total

profit earned on average by explorers, imitators and the Platform firm are displayed side by side. Some considerations worth of highlighting are, first, that explorers' profit is never significantly affected by the presence of a Platform firm. As the market discovers products that are better suited to consumers' tastes, the value of completely random discoveries by explorers falls down. The latter holds true in all the simulated scenarios and is affected by the presence of a Platform firm only marginally. Second, the Platform firm subtracts market shares mostly from imitators, almost on a one-to-one basis. However, even with its more efficient imitation algorithm, the Platform firm is rarely able to gain a very large share of the imitators' market, as it on average manages to produce profit, in the very last simulated periods, that is about 12% to 22% of imitators' profit, and about 10% to 18% of total profit earned by imitators and explorers taken together. Combined with the observation that across all simulations, the Platform firm may or not land in the first market positions based on the specific dynamic path generated by the interaction between consumers and new products, this set of results casts doubts on the prior, according to which insider imitation by the Platform firm may lead to larger market concentration and a leadership position for the Platform firm. By the same token (compare to previous Fig. 2), although in the scenario with many visible products the Platform firm reaps larger profit away from imitators, this does not imply a loss in welfare. The reason lies in the ability for the better imitation algorithm by the Platform firm to discover product characteristics that are more appealing to consumers which, consequently, raise the average utility obtained per purchase.

As a concluding remark, I may note that my simulation results are compatible with empirical results presented in Crawford et al. (2022). In their sample of products sold on the Amazon marketplace, Crawford et al. (2022) find that almost 50% of the revenue generated by Amazon products comes from products that no other merchant previously offered on the marketplace, which is compatible with the argument that platform providers engage in partial imitation, rather than in copycat imitation where they would more directly steal market shares from an existing product. Moreover, Crawford et al. (2022) find that "Amazon entry is associated with mild market expansion, with no evidence of business stealing" both measured against direct substitutes or against close substitutes, and both looking at total markets including Amazon sales or at sales from third party traders only. The latter result, too, coincides with simulations where the presence of a Platform firm produces an increase in overall trading volumes, although in my simulations the Platform firm does steal market shares from imitators (though the latter manage to keep the majority of market shares for themselves).

4 Conclusions

In this paper I study insider imitation in an online market with product differentiation and partial imitation. I showed how the presence of an imitator, who may benefit from accessing market-wide data not available to competitors, may speed up the rate of innovation via imitations, without necessarily leading to an increase in market concentration. Consumers on average may benefit from insider imitation as this improves the average utility they obtain from their purchases. Competing firms who also engage in

partial imitation benefit and lose simultaneously: on the one hand, part of their market shares are won by the platform provider, however on the other hand imitators can imitate the imitator, thus improving their products at a faster rate than otherwise. The net overall effects of such dynamic simultaneous effects on market outcomes depends on a number of parameters representing specific market characteristics, controlling for which I may claim that the most likely effect of insider imitation of the kind examined here is to increase welfare and reduce market concentration.

The analysis illustrated so far casts doubts on the latest stance of the European Commission's antitrust authority as an effective approach to insider imitation on online marketplaces. I have shown that, under a large range of model parametrizations, the presence of an insider imitator who uses unpublished data to engage in partial imitation is likely improving welfare or at least, not reducing it. Moreover, in most of the scenarios, the presence of such insider imitation reduces market concentration. Forbidding the use of nonpublic data by Amazon, as the recent decision of the Commission implies (see the decision 2023/C 87/0, in particular related to case AT.40462), may have negative consequences for consumers and might achieve a suboptimal level of efficiency in market competition. This claim adds up to a number of other studies, mentioned in the Introduction, which also bring a cautionary word of advice by showing that the causal link between insider imitation, market efficiency and policy actions is, in the case of online platforms, a complex one.

My analysis has some important limitations which will hopefully be addressed by future research. The Platform firm in the simulations does not exploit other means that are generally available to a platform provider, like for instance the ability to manipulate the way search results are shown to users (self-preferencing), or the ability to offer complementary services only for selected products (e.g., the Amazon Prime program). Thus, my results cannot completely rule out that a platform provider engaging in the type of insider imitation here discussed could displace competitors and reduce welfare. My claim is that the sole existence of insider imitation by means of data exploitation is not, *per se*, reason for concern and, moreover, policy action may end up reducing consumer's welfare. Another point deserving further scrutiny is the possibility that competitors reduce their rate of innovation as the Platform firm acquires market shares. The available empirical evidence on the Amazon marketplace suggests, indirectly, that this should not be a major issue as imitation does not happen immediately (thus granting innovative products some lead time to reap returns) and, at least in the case of Amazon, the platform provider tends to acquire minority shares of such markets. An extended model that explicitly takes into account investment decisions might shed light on the interaction between these different dimensions. Finally, as stated already the utility function that was assumed to represent consumer preferences is independent from the variety of products available on the market. An amended model might explore a different set of preferences and quantitatively estimate whether reductions in product variety may welfare-dominate the other effects studied here.

Appendix A

Additional figures.

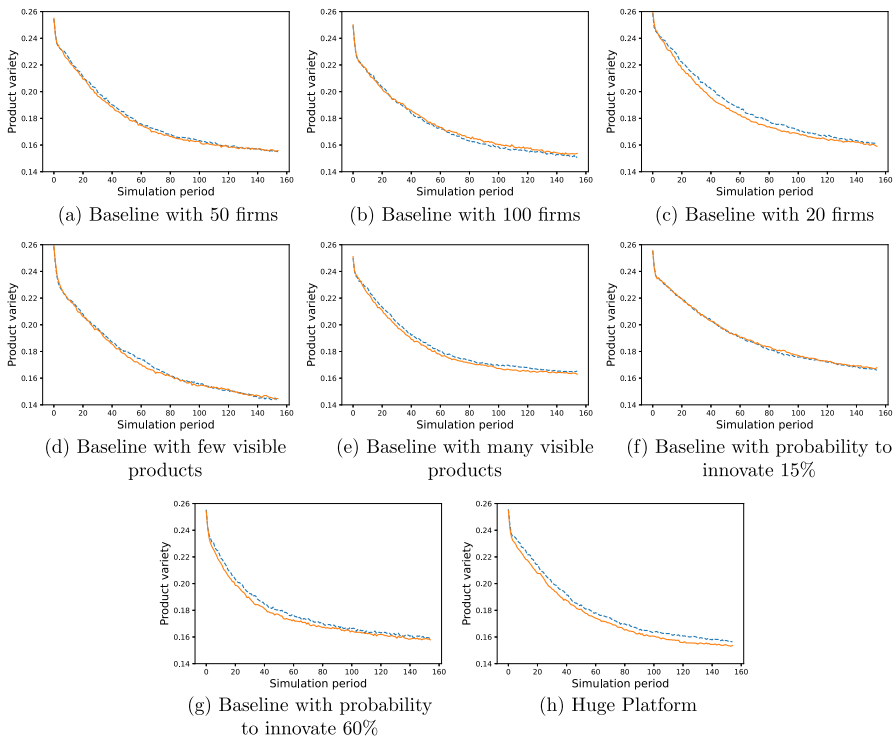


Fig. 5 Product variety (average across simulated scenarios) - With a Platform firm (orange continuous lines) and without it (dashed blue lines)

Appendix B

This Appendix reports the results of a search on the Amazon.com Web site for the string of text “cloth hanger”. This search was performed on 8 November 2023 from a position based in Germany (I mention this, as it might be the case that the Amazon.com Web site discriminates products based on the IP address of the user). I only used the internal search engine of Amazon.com.

Figure 8 shows the products offered by Amazon with its own brand, Amazon Basics. The first of these three products is made of stainless steel and comes in two colors, the second is made of wood, while the third is made of plastic and is offered in ten different colors (the color is made different for the curved element and for the metal hook).

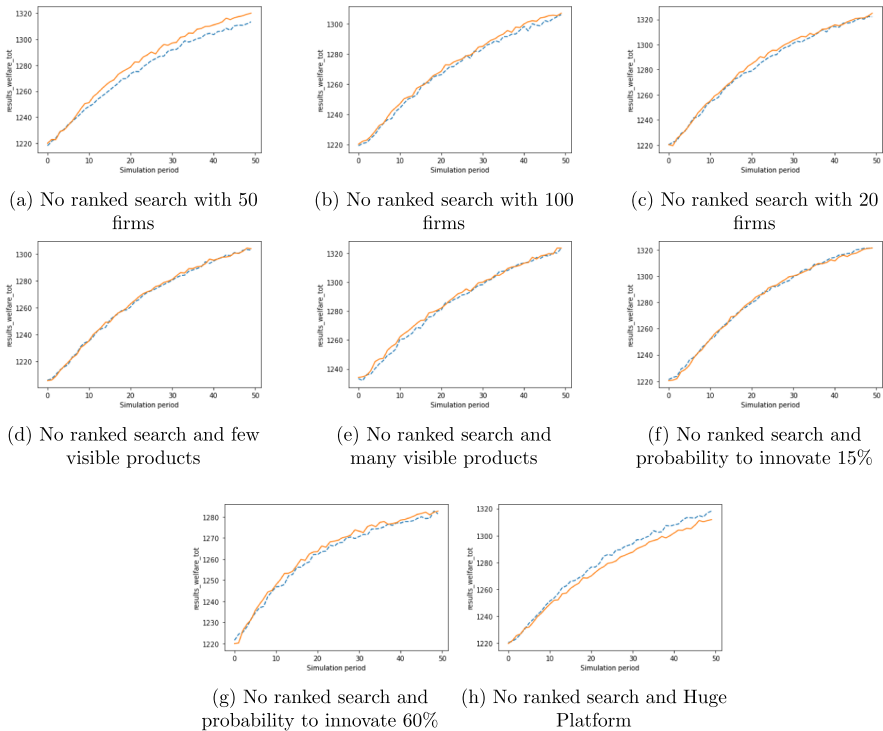


Fig. 6 Total welfare (average across simulated scenarios), alternative scenarios without ranked search - With a Platform firm (orange continuous lines) and without it (dashed blue lines)

Compared to all other products found on the same platform, a representative sample of which is displayed in Figs. 10 and 11 (these, together with those shown in Fig. 9, are the 24 top-rated products from the first two pages returned by the Amazon.com search engine), one may immediately see that the Amazon-brand products offer a unique bundle of features. Referring to the metal-made Amazon Basics hanger only, it features: a metal structure with anti-slippery coating; swivel hook; two clips with anti-slippery rubber; indenting on the sides of the curved element to allow for the hanging of thin shoulder straps. While all of these features may be individually found on some competing products, this specific bundle is unique to Amazon's. Other products focus more on aesthetics, price or on different functionalities. The most similar competing products are: the second-last depicted at the end of Fig. 11 (which however, differently from Amazon's own variety, has the curved element made of plastic); and the last one (which has a different design for the clips with a silicon coating and the use of biodegradable plastic). Figure 9 reports the three best-selling products according to information returned by the Amazon.com search engine which displays an estimate for the number of sales in previous month. None of them comes from an Amazon brand.

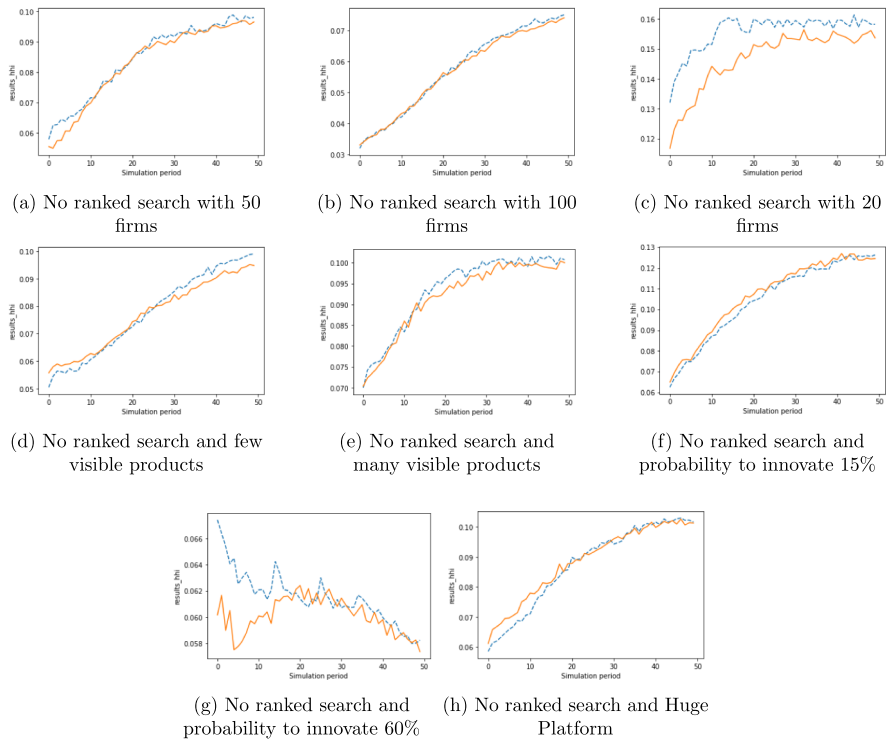


Fig. 7 Market concentration - Herfindahl-Hirschman Index (HHI, average across simulated scenarios), alternative scenarios without ranked search - With a Platform firm (orange continuous lines) and without it (dashed blue lines)



Fig. 8 Search results on Amazon.com for “cloth hanger” - Amazon own brand products



(a) 20k sold in previous month (b) 10k sold in previous month (c) 7k sold in previous month

Fig. 9 Best-selling products in previous month (estimated No. of sales returned by Amazon.com)



Fig. 10 Search results on Amazon.com for “cloth hanger” - selected results based on reported No. of reviews on Amazon.com, Part 1



Fig. 11 Search results on Amazon.com for “cloth hanger” - selected results based on reported No. of reviews on Amazon.com, Part 2

Acknowledgements I would like to thank the Editor and two anonymous Referees of the Journal of Evolutionary Economics for very thorough commentary which greatly helped improving the paper. My gratitude goes also to the participants to the 27th Annual Workshop on Economics with Heterogeneous Interacting Agents (WEHIA), held in July 2024 in Bamberg, Germany, for their useful comments and suggestions. All remaining errors are the sole responsibility of the author.

Author Contributions D. d’A. wrote the entire manuscript and the simulation code in Python, ran all simulations and prepared figures and tables summarizing the main simulation results.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Ethical Approval Not applicable.

Conflict of interest The authors declare no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Cantner U (2007) Firms' differential innovative success and market dynamics. *Jena Econ Res Pap* 2007–078
- Cantner U, Hanusch H (2002) Evolutionary economics, its basic concepts and methods. *Editing Economics: Essays in Honour of Mark Perlman* 1991:182
- Cantner U, Savin I, Vannuccini S (2019) Replicator dynamics in value chains: Explaining some puzzles of market selection. *Ind Corp Chang* 28(3):589–611
- Choi JP, Kim K, Mukherjee A (2023) sherlocking'and information design by hybrid platforms. Available at SSRN 4332558
- Crawford GS, Courthoud M, Seibel R, Zuzek S (2022) Amazon entry on amazon marketplace. *CEPR Discussion Papers* (17531)
- d'Andria D, Savin I (2018) A win-win-win? motivating innovation in a knowledge economy with tax incentives. *Technol Forecast Soc Chang* 127:38–56
- Dawid H (2006) Agent-based models of innovation and technological change. *Handbook of Computational Economics* 2:1235–1272
- Etro F (2021) Product selection in online marketplaces. *J Econ Manag Strateg* 30(3):614–637
- Etro F (2022) The economics of amazon. Available at SSRN 4307213
- European Commission (2022) Antitrust: Commission seeks feedback on commitments offered by amazon concerning marketplace seller data and access to buy box and prime. https://ec.europa.eu/commission/presscorner/detail/en/IP_22_4522. Accessed 11 Aug 2023
- Faherty E, Huang K, Land R (2017) The amazon monopoly: Is amazon's private label business the tipping point? MPRA Paper, Munich Personal RePEc Archive
- Gutierrez G (2021) The welfare consequences of regulating mazon. Available at SSRN 3965566
- Hagiu A, Teh T, Wright J (2022) Should platforms be allowed to sell on their own marketplaces? *Rand J Econ* 53(2):297–327
- Khan LM (2016) Amazon's antitrust paradox. *Yale LJ* 126:710
- Kiesling E, Günther M, Stummer C, Wakolbinger LM (2012) Agent-based simulation of innovation diffusion: a review. *CEJOR* 20:183–230
- Kim L, Nelson RR (2000) Technology, learning, and innovation: Experiences of newly industrializing economies. Cambridge University Press
- Ma T, Nakamori Y (2005) Agent-based modeling on technological innovation as an evolutionary process. *Eur J Oper Res* 166(3):741–755
- Madsen E, Vellodi N (2024) Insider imitation. *J Polit Econ* (forthcoming)
- Nelson RR, Winter SG (2009) An evolutionary theory of economic change. Harvard University Press
- Rochet J, Tirole J (2006) Two-sided markets: a progress report. *Rand J Econ* 37(3):645–667
- Tesfatsion L (2003) Agent-based computational economics: modeling economies as complex adaptive systems. *Inf Sci* 149(4):262–268
- Tesfatsion L, Judd KL (2006) Handbook of computational economics: agent-based computational economics. Elsevier
- Vellodi N (2018) Ratings design and barriers to entry. *NET Institute Working Paper No. 18-13*
- Wang L, Wu B, Pechmann C, Wang Y (2023) The performance effects of creative imitation on original products: Evidence from lab and field experiments. *Strateg Manag J* 44(1):171–196
- Yildizoglu M (2002) Competing r&d strategies in an evolutionary industry model. *Comput Econ* 19:51–65

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.