ARTICLE IN PRESS

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/eor

Analogy-based life cycle forecasts with pre-release buzz

Oliver Schaer [a,*], Nikolaos Kourentzes [b], Robert Fildes [c]

[a] LeBow College of Business, Drexel University, USA
[b] School of Informatics, University of Skövde, Sweden
[c] Centre for Marketing Analytics and Forecasting, Lancaster University Management School, UK

ARTICLE INFO

Keywords:
Forecasting
OR in marketing
Google trends
Pre-launch forecasting
New product adoption

ABSTRACT

With shorter product life cycles and increased competition, forecasting new product sales prior to launch is vital. We contribute to the new product forecasting literature by augmenting analogy-based approaches with pre-release online search traffic. In contrast to existing research, which relies solely on pre-release buzz during the launch phase, we consider life cycle sales. We propose a model of pre-release online search traffic and market potential, connecting the two to support pre-launch decision-making. We validate this relationship with an empirical experiment on sequential video game sales. Our findings support that pre-release online search traffic contains predictive information up to 18 weeks before release and can increase life cycle sales forecast accuracy by up to 21 %. The explanatory power of pre-release online search traffic varies across product generations. This evolution opens up marketing opportunities and highlights the importance of managing pre-release search interest. Our approach can be implemented with minimal data requirements, making it a versatile and accessible tool for firms. We provide extensive managerial findings and a way forward for incorporating this approach into "new to the world" products.

1. Introduction

With shorter product life cycles and increased competition, generating accurate pre-launch forecasts is vital for companies to plan their marketing expenditures and production effectively. Companies must accurately predict the adoption shape and overall lifetime sales (i.e., market potential) to ensure a product's success (Goodwin et al., 2014). Since many new products represent incremental innovations (Markham & Lee, 2013), decision-makers can often derive the adoption shape through analogies of previous or similar products (e.g. Hu et al., 2019; Lenk & Rao, 1990), but struggle to predict the new product's overall market potential (Trusov et al., 2013). These issues are particularly relevant to the video game industry, where market potential is a key parameter for publishers and developers in planning their operational and marketing rollouts.

Recent literature has identified major weaknesses in the primary methods used to estimate overall market potential. Traditionally, practitioners estimate the market potential of new products through expert judgment (e.g., Bass et al., 2001; Kim et al., 2013; Mahajan et al., 1986). However, substantial evidence indicates experts are biased when forecasting the success of a new product (Belvedere & Goodwin, 2017; Feiler & Tong, 2022; Markovitch et al., 2014; Tyebjee, 1987). Alternative approaches to predicting overall market potential include consumer surveys (e.g., Chintagunta & Lee, 2012; Eliashberg et al., 2000; Moon et al., 2016) or conjoint analyses (e.g., Lee et al., 2006; Orbach & Fruchter, 2011). However, these methods are costly, and consumer preferences often change during the pre-release phase, invalidating survey results (Meeran et al., 2017). Therefore, practitioners require more efficient and effective ways to identify their market potential for incremental product innovation and brand extensions.

Research on pre-release buzz (PRB) suggests that it has the potential to improve forecasts; however, scholars have yet to consider its use when paired with analogies. PRB is an aggregate measure of consumers' anticipation for a new product (Houston et al., 2018; Mandler et al., 2025), like search traffic, blogs, and social media. Forecasting studies that use PRB, exclusively focus on predicting the initial product release period, i.e., the first few weeks after launch. Therefore, the suitability of PRB for long-term forecasts remains an open question of managerial importance, such as for planning short-life cycle products with few reorder points (Mostard et al., 2011). Moreover, these studies are based on data samples that require companies to have access to competitor sales data. This defeats the idea that PRB is a cost-effective and competitive intelligence resource (Schaer et al., 2022).

To this end, we propose a more parsimonious and data-economic method by integrating analogy-based life cycle forecasting with the PRB literature. More specifically, we leverage analogies from previous prod-

* Corresponding author.
E-mail addresses: info@oliverschaer.ch (O. Schaer), nikolaos@kourentzes.com (N. Kourentzes), r.fildes@lancaster.ac.uk (R. Fildes).

https://doi.org/10.1016/j.ejor.2025.12.003
Received 31 January 2025; Accepted 1 December 2025
Available online 7 December 2025

Please cite this article as: Oliver Schaer et al., *European Journal of Operational Research,* https://doi.org/10.1016/j.ejor.2025.12.003

uct generations with Pre-Release Online Search Traffic (PROST) from Google Trends. Search traffic is particularly relevant as a proxy for consumer interest in a product (Du & Kamakura, 2012; Houston et al., 2018; Stephen & Galak, 2012) and for marketing expenditures (Hu et al., 2014; Srinivasan et al., 2016). Contrary to prior work, the proposed model utilises only firm-owned and publicly available PROST data, eliminating the need for competitors' sales data and simplifying its use and implementation. In this study, we (i) test the relation between PROST and the market potential of a new product generation, (ii) investigate the leading properties of PROST, and (iii) analyse how the explanatory power of PROST changes as product generations mature. Our empirical case study of forecasting pre-launch video game sales demonstrates that PROST is a valuable source for predicting both short- and long-term sales, requiring only two pre-release periods. The PROST model demonstrates higher explanatory power for overall market potential than first-week sales, achieving up to a 21 % increase in life cycle sales forecast accuracy compared to past market potential. Companies have employed various strategies to spark pre-release interest in video games, alternating the diffusion path and final market potential. For example, Rockstar turned a data leak of the highly anticipated "GTA6" game into increased visibility, with the official trailer achieving record-breaking views (Loveridge, 2025). Blizzard established a large fanbase before releasing "Overwatch" by investing heavily in pre-release advertising and strategically managing beta access (Tassi, 2016). Similarly, the developer of the indie game "Animal Well" created deep engagement by embedding hidden puzzles across the game's promotional channels, building an active Discord community that helped to sell over 300,000 copies in just 9 months (Machkovech, 2024). By leveraging such public information, our approach can provide valuable predictive inputs to decision-makers tasked with decisions related to product pre-launch, launch, and post-launch phases, complementing the existing literature, which is primarily focused on launch-phase sales.

Section 2 reviews the relevant literature on PRB. Section 3 elaborates how search traffic can be modelled to estimate the market potential to augment analogies. Section 4 empirically demonstrates the value of the proposed market potential estimation using various life cycle curves. This provides new insights into how the predictive value of pre-release online search traffic evolves. These are discussed in Section 5 and include practical implications for the video game industry regarding pricing, return on investment, sourcing strategy, release timing, and ad spending. Lastly, we provide the reader with a modification to forecast first-generation products, enabling decision-makers to use PROST across the complete life cycle of multi-generation products.

2. Forecasting with pre-release buzz

To understand how PRB can augment analogy-based forecasting, we first examine the prior literature on its predictive capabilities. This stream of literature heavily centres around the entertainment industry with studies forecasting box office sales, video games, and music album sales that utilise a variety of PRB sources, including blogs (e.g., Kim & Hanssens, 2017; Kim et al., 2015), microblogs (e.g., Asur & Huberman, 2010; Lehrer & Xie, 2021), forums (e.g., Craig et al., 2015; Wang et al., 2010), and search engines (e.g., Kulkarni et al., 2012; Xiong & Bharadwaj, 2014). The literature also reports applications in the apparel industry, where researchers leverage colour trends from Google or Twitter mentions to predict the next season's sales (Fu & Fisher, 2023; Skenderi et al., 2024). In the following, we discuss three key areas in which we identify significant limitations in the current literature that inform our research questions and provide a summary to support our research positioning.

2.1. Short-term forecasting focus in PRB research

Prior research demonstrates that PRB improves accuracy during the launch phase, but provides limited guidance for predictions beyond one month. Only two studies on fashion retail forecast the entire full-price season, spanning 6 to 12 weeks (Fu & Fisher, 2023; Skenderi et al., 2024). Researchers argue that sales patterns can be determined from the initial week (e.g., movies, Kulkarni et al., 2012), forcing decision-makers to switch to an alternative forecasting model. However, in practice, switching models may be undesirable, as it can introduce inconsistencies between forecasts, leading to different forecast outlooks across models and similar differences in the resulting decisions. Moreover, this involves additional training and validation requirements, as models for different targets and horizons need to be built. Any short-term forecast model requires additional data points before it can produce reasonable forecasts (see Lei et al., 2023, for an example of switching models after 3 months).

The focus on short-term predictions reflects the assumption that PRB's predictive value diminishes after launch. As a result, most studies use regression or machine learning to estimate the point forecasts, with PRB serving as an exogenous indicator and sales as their direct target. A notable exception is Wang et al. (2010), who use PRB to improve the estimation of diffusion curves. However, their approach focuses on updating only the parameters that control the shape of adoption. This raises the question of how useful PRB is with respect to the market potential parameter that controls overall adoption, as shown in several post-release data studies embracing advertising activities (e.g., Toubia et al., 2014; van der Lans et al., 2010) or online reviews (Dellarocas et al., 2007).

Yet, across the surveyed studies on forecasting with PRB, the connection between PRB and total life cycle sales has been neglected and remains an open question of managerial importance, particularly for planning short-life cycle products with few reorder points. This lack of focus on long-term demand directly motivates our first hypothesis, which is whether PRB is a useful predictor of total adoption, effectively extending its use beyond the initial launch phase.

Hypothesis 1. *Pre-release buzz (PRB) improves forecast accuracy for the market potential (i.e., total life cycle sales) beyond the initial week of sales.*

2.2. Lead time dynamics of PRB

The practical importance of forecasting in the life cycle, often with long lead times, is evident. Yet, few studies have systematically examined the predictive ability of PRB for different lead times. Kim and Hanssens (2017) note that most of the literature on box-office forecasting focuses on predicting near release day, leading to a nowcast that limits managerial relevance and model validation. The longest lead time in the reviewed literature extends 151 days before the product launch (Xiong & Bharadwaj, 2014). Even at this relatively early stage, their forecasts outperformed benchmarks based on static product characteristics, ad spending, or pre-orders. This suggests that the dynamics of user-generated content may offer greater explanatory value than company-internal data.

However, studies disagree on what lead-time PRB provides the highest predictive value. While several studies suggest that buzz appearing closer to the product launch is most useful (Chintagunta & Lee, 2012; Goel et al., 2010; Wang et al., 2010; Xiong & Bharadwaj, 2014), Kulkarni et al. (2012) report that forecasts of box-office sales made four weeks before launch are more accurate than those made the week prior to launch. Even more remarkably, Fu and Fisher (2023) finds that the predictive performance remains consistent for all lead times up to 22 weeks. Additionally, Goel et al. (2010) shows diminishing benefits when other key indicators, such as past sales performance or the production budget, became available. This aligns with research on short-term forecasting with online buzz for established products, where user-generated content is less effective as a predictor once the product has been on the market (Fu & Fisher, 2023; Marchand et al., 2017).

These conflicting findings indicate that the relationship between lead time and predictive accuracy may vary systematically across different forecasting contexts. Understanding these dynamics becomes particularly important for strategic decisions regarding return on investment, replenishment cycles, or infrastructure utilization, where managers need reliable forecasts well in advance of product launch. Similarly, it is unclear from these studies whether these dynamics differ between short-term and long-term forecast scenarios. The inconsistent evidence regarding lead-time dynamics, combined with the lack of systematic investigation across different forecast horizons, motivates our second hypothesis about how PRB's predictive value evolves across lead times.

Hypothesis 2. *Pre-release buzz (PRB) improves the accuracy of sales forecasts beyond the initial sales week, even for longer lead times.*

2.3. Sequential adoption dynamics and franchise maturity

These days, a large portion, if not the majority, of new entertainment and short-life cycle products represent a sequel (Hennig-Thurau & Houston, 2019) or incremental innovation (Markham & Lee, 2013). However, our literature review reveals that only a few studies on PRB forecasting specifically account for sequels in their models, for example, by coding them as dummy variables (also noted in Marchand, 2016). In other words, a majority of studies treat new products independently, regardless of their potential close connection to a previous product.

The limited studies that do examine sequels of entertainment products indicate that such products tend to follow a steeper trend (Xiong & Bharadwaj, 2014) and start with a higher initial PRB level (Craig et al., 2015; Xiong & Bharadwaj, 2014) than prior releases. Goel et al. (2010) considered the lifetime revenue of predecessors as a direct predictor, which had higher explanatory power than search traffic alone. Similarly, although in the context of fashion retail, Fu and Fisher (2023) builds its base model on past seasons and previous years' fashion sales. These findings highlight the importance of modeling new products within their sequential context rather than as isolated entities.

Although there is substantial work in the area of new product forecasting with analogies that rely on previous product generation information (see, Goodwin et al., 2014), none of the reviewed studies specifically investigated how the predictive ability of PRB changes across product generations as the franchise matures and consumers have increased product awareness. This presents a missed opportunity, given companies' heavy reliance on sequential release, and highlights the need for a better understanding of how generational dynamics affect forecast performance. As franchises mature, several factors may influence the effectiveness of PRB, including consumer familiarity and established fan bases, which affect interaction with pre-release content and PRB measurement.

Understanding these dynamics becomes particularly important for companies managing sequential product portfolios, as it enables them to effectively adjust their pre-release marketing strategies across different product generations. We therefore formulate a third hypothesis that examines whether the efficiency of pre-release buzz varies across product generations.

Hypothesis 3. *The predictive value of pre-release buzz (PRB) varies systematically as the franchise matures.*

2.4. Literature summary and research positioning

Table 1 summarises the reviewed literature, highlighting the patterns that emerge by buzz predictor, lead time, forecast horizon, target variable, sample size, estimation approach, and whether the estimation includes information from previous generations. We also indicate whether the analysis is based on in-house sales data, i.e., organisations can run the analysis without purchasing competitor sales information.

Our literature review has identified three interconnected research questions that, when addressed together, provide a clearer understanding of PRB's predictive capabilities for pre-launch life cycle forecasts. First, the predominant focus on the launch period overlooks the critical need for life cycle predictions that support strategic decisions, like pricing, capacity planning, and marketing allocation. Second, many of these decisions require a substantial lead time. However, since studies indicate that PRB effectiveness varies with lead time in short-term forecasting, it remains unclear how this affects life cycle forecasts and what the practical limits of PRB are. Third, very few studies examine how PRB's predictive power changes across product generations, offering a research opportunity to leverage forecasts through analogies.

Except for examples from the fashion industry (Fu & Fisher, 2023; Skenderi et al., 2024), it is important to note that all the surveyed literature bases its estimations on samples that include multiple producers or brands. Arguably, this facilitates the study of overall market dynamics and the use of more complex models. However, this does not reflect reality for many companies, as they often have access to only a limited number of in-house products and, based on the approaches available in the literature, are likely required to rely on some form of competitive intelligence to train their model successfully (see Schaer et al., 2022). Therefore, the goal of this research is to develop a more parsimonious forecasting approach that relies solely on publicly available PRB and past in-house product sales information, making it more practical to implement while addressing the three hypotheses outlined above.

3. Modelling sequential products with analogies and pre-release buzz

To evaluate our hypotheses, we combine two research streams: analogies and PRB. One common way to predict demand for a new product pre-release is to forecast based on analogies from previous generations or similar products (e.g., Baardman et al., 2018; Hu et al., 2019; Van Steenbergen & Mes, 2020). Following this model route enables us to develop a more parsimonious approach that addresses the aforementioned product data scarcity faced by many companies. We test whether it is beneficial to supplement past adoption profiles with PRB to estimate market potential. In this section, we formulate the connection between overall product sales (i.e., analogies) and the pre-launch buzz of product generations.

We define market potential as the overall life cycle sales (cumulative adoption), corresponding to parameter m in Bass (1969). The cumulative adoption (A_t) is:

$$A_t = mF_t(\cdot), \tag{1}$$

where m is the market potential, and $F(\cdot)$ can be any function defining the cumulative shape of adoption across time (we discuss life cycle functions in Section 4.2). This approach requires estimating a single model across the entire life cycle, unlike methods that separately estimate first-week and post-release sales.

Our approach also makes a direct contribution to the literature on diffusion modelling. Many studies focus on detecting the adoption shape and omit estimating market potential, assuming it to be known (e.g., Goodwin et al., 2013; Lee et al., 2014). In the case of Hu et al. (2019), the demand planner estimates the market potential, which is common in practice and frequently done through judgment adjustments (Goodwin et al., 2014) or by relying on demographic data (e.g., Talukdar et al., 2002). As such, we aim to establish an approach that reflects the current market potential at a much lower cost than traditional market intelligence.

We focus on sequential product launches, which provide natural analogies for forecasting. Today, companies predominantly invest in incremental innovations (Markham & Lee, 2013), or brand extensions, particularly in entertainment products (Hennig-Thurau et al., 2009; Sood & Drèze, 2006). Regardless of the specific context, there is a strong link between the previous and new product generations, as they share similar product features. Without this distinctive connection, one must gather additional information about related products and mar-

Table 1
Forecasting with pre-launch buzz.

Study	Buzz pred.*	Max. lead†	Max. horizon†	Target var.⋄	# series	Model⋆	Sequel	Sales source
This study	GT	26 w	52 w	VG	251	AY	sales	in-house
Skenderi et al. (2024)	GT	24 w	6 w	FN	5577	ML	–	in-house
Xiong and Bharadwaj (2014)	GT;BL;FM	22 w	3 w	VG	681	FR	binary	x-brands
Fu and Fisher (2023)	GT;TW	18 w	12 w	FN	1313	ML	sales	in-house
Goel et al. (2010)	GT	6 w	1 w	BO	119	LR	–	x-brands
	GT	6 w	1 w	MA	307	LR	–	x-brands
	GT	6 w	1 m	VG	106	LR	sales	x-brands
Kulkarni et al. (2012)	GT	4 w	1 w	BO	61	LR	–	x-brands
Kim and Hanssens (2017)	GT;BL	3 w	1 w	BO	137	LR	–	x-brands
Kim et al. (2015)	BL	1 w	1 w	BO	212	ML	–	x-brands
Divakaran et al. (2017)	BL	1 w	1 w	BO	373	PR	–	x-brands
Ding et al. (2017)	FB	1 w	1 w	BO	64	LR	–	x-brands
Wang et al. (2010)	FM	1 w	2 w	BO	51	DC	–	x-brands
Craig et al. (2015)	FM	3 w	1 w	BO	62	LR	binary	x-brands
Lehrer and Xie (2021)	TW	1 d	1 d	BO	178	ML	–	x-brands
Asur and Huberman (2010)	TW	1 d	1 d	BO	24	LR	–	x-brands

∗ BL = Blog, FB = Facebook, FM = Forum, GT = Google Trends, TW = Twitter;
† d = day(s), w = week(s), m = month(s);
⋄ BO = Box Office, FN = Fashion, MA = Music Albums, VG = Video Games;
⋆AY = Analogy, DC = Diffusion curve, FR = Functional Regression, LR = Linear Regression, ML = Machine Learning.

Table 2
Market potential estimation with PROST.

Model	Equation	Required # coef.	Earliest estimation	Model description
M.1	$\hat{m}_j = c_0 \cdot \Delta\text{PROST}_j^{c_1} \cdot m_{j-1}^{c_2} \cdot \varepsilon_j$	3	$j = 5$	Full model with autoregressive term
M.2	$\hat{m}_j = c_0 \cdot \Delta\text{PROST}_j^{c_1} \cdot \varepsilon_j$	2	$j = 4$	Removed autogregressive term
M.3	$\hat{m}_j = c_0 \cdot \Delta\text{PROST}_j^{c_1} \cdot m_{j-1} \cdot \varepsilon_j$	2	$j = 4$	Fixed autoregressive coefficient ($c_2 = 1$)
M.4	$\hat{m}_j = \Delta\text{PROST}_j^{c_1} \cdot m_{j-1} \cdot \varepsilon_j$	1	$j = 3$	Removed intercept term
M.5	$\hat{m}_j = \Delta\text{PROST}_j \cdot m_{j-1} \cdot \varepsilon_j$	0	$j = 2$	PROST unit change only
M.6	$\hat{m}_j = \Delta\text{PROST}_j^{\frac{1}{2}} \cdot m_{j-1} \cdot \varepsilon_j$	0	$j = 2$	damped PROST unit change only

kets to identify useful analogies (Goodwin et al., 2014). We provide an extension to the proposed model that looks sideways for forecasting the first generation without direct predecessors in Section 5.1.

Studies comparing PRB sources show mixed results: While Xiong and Bharadwaj (2014) found blog posts more accurate than search data, Kim and Hanssens (2017) achieved better forecast accuracy using search alone. Studies have also compared context- and volume-based measures, overall suggesting volume provides more predictive power (Lehrer & Xie, 2021). For this reason, this research will focus on the volume of online search traffic. Typically, search traffic is represented as a relative measure, reflecting interest in a product (Chumnumpan & Shi, 2019; Du & Kamakura, 2012; Houston et al., 2018; Stephen & Galak, 2012). We have investigated both additive and multiplicative formulations and found the latter superior.

Table 2 introduces the multiplicative models for estimating the market potential of the jth generation. In the full form, model M.1 is an autoregressive indicator with differenced (Δ)PROST between the two consecutive generations as an explanatory variable with the error assumed to be a log-normal $\varepsilon_j \sim \log \mathcal{N}\left(0, \sigma_j^2\right)$. Moreover, the autoregressive process accounts for carryover between product generations (Situmeang et al., 2014). However, as indicated in the second-to-last column, the full model requires at least 5 product generations to estimate the required number of parameters. For this reason, we introduce alternative, restricted versions: M.2 is without the autoregressive term $m_{j-1}^{c_2}$, and M.3 incorporates past market potential with a fixed coefficient of $c_2 = 1$. M.4 and M.5 are further restricted by dropping PROST's intercept c_0 and coefficient estimation c_1.

Lastly, we introduce M.6, which modifies M.4 and M.5 by dampening the effect of PROST. Reducing the PROST coefficient towards zero underlies the principal idea of shrinkage (Green & Armstrong, 2015; Hastie et al., 2015) and represents a more conservative approach. Like M.5, M.6 employs a fixed coefficient, which can be advantageous when the sample size is limited; in both cases, only two observations are required for estimation.

This study differentiates itself from the literature on online buzz in the following ways: First, our approach is very economical in requiring only a single franchise, in contrast to the majority of studies that used large data samples consisting of multiple publishers (e.g., Goel et al., 2010; Lehrer & Xie, 2021; Xiong & Bharadwaj, 2014). This also allows us to observe dynamics across generations directly. Second, many studies focus on sales during the launch period. We provide holistic life cycle forecasts, giving richer information to decision-makers (see max. horizon in Table 1). Third, studies that use diffusion curves either focus on updating adoption parameters rather than market potential (Wang et al., 2010) or are conducted within a nowcasting framework post-release (Chumnumpan & Shi, 2019). We focus on market potential, a key quantity that informs many pre-launch planning decisions.

4. Pre-release video game adoption modelling

We evaluate our approach using video game sales data. With sequels dominating US sales (ESA, 2023)[9 of 10 bestsellers in 2022] and active online gaming communities, this sector provides an ideal testing case. We utilise physical video games sales data from VGChartz (http://www.vgchartz.com), a data source also used in other pre-release research studies (e.g. Marchand et al., 2017; Xiong & Bharadwaj, 2014). Our dataset comprises weekly sales for 251 games across 57 popular franchises such as *Call of Duty* and *FIFA* from multiple publishers, with 2 to 11 sequential product instalments (average 4.4) per franchise.

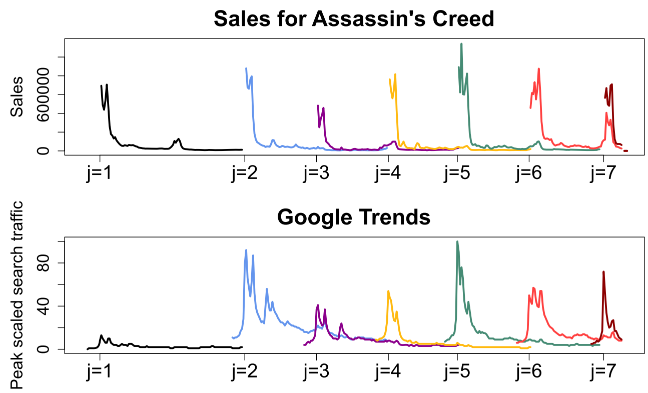


Fig. 1. Generations of the Assassin's Creed compared to Google Trends.

In this research, we use PROST from Google Trends as a proxy for PRB. Fig. 1 illustrates the relationship between observed sales and PRB, where the correlation between the two is apparent, motivating the models introduced in Table 2. Situmeang et al. (2014) find a significant carry-over effect between video game generations, i.e., a generation with good or bad user experience will impact the upcoming release, further strengthening the specified models.

4.1. Pre-release online search traffic (PROST) data

To represent PROST, we collected Google Trends time series for each game. Keywords were selected using Google Knowledge Graph entities to ensure robust search coverage, which accounts for spelling errors and regional spelling variations, while also filtering out traffic unrelated to the particular game (e.g., Schaer et al., 2019), defaulting to the game title when unavailable. For modelling, we consider two parameters: the window length (w) of the recorded search volume and the lead time (l) before the launch date T. Consequently, for these windows, we then sum Google Trends (GT) so that we have one data point of PROST for each generation j:

$$\mathrm{PROST}_j = \sum_{k=1}^{w} \mathrm{GT}_{j,(T-l-k)}. \tag{2}$$

In the experimental setting, it is possible to either expand the PROST window over time or keep it fixed. While Xiong and Bharadwaj (2014) report gradually improved forecast accuracy towards the release date as they expand their window size, Kulkarni et al. (2012) find better performance in smaller window sizes further away from the release. For a fixed window of size one, Kim and Hanssens (2017) find better performance closer to release. Since it is unclear whether changes in forecast accuracy are due to the window length or the lead time, we investigate both dimensions.

Google Trends is peak-scaled (0-100), which means different data requests are inherently non-comparable and limited to 5 keywords. We addressed this limitation by using a neutral scaling keyword ("marker") to achieve comparable data across all games, similarly to Schreiner et al. (2025). We scaled each video game's popularity relative to the reference keyword by calculating the sum of the scaling keyword's popularity for a specific period and scaling it to the sum of the video game keyword (see, Kim & Hanssens, 2017; Schaer et al., 2019, for a discussion on scaling and data reliability of Google Trends).

We ran the experiment with a preset rolling window evaluation from 26 weeks to 1 week prior to release, and then reran it with different PROST window sizes ranging from 4 to 12 weeks. At a maximum, we considered search traffic information up to 38 weeks prior to release, an approach similar to that of Xiong and Bharadwaj (2014), which included PRB 26 weeks (or 181 days) prior to launch.

Table 3
Life cycle curves.

Bass	$\hat{F}(j,t) = \left(1 - \exp\left(-(\hat{p}_j + \hat{q}_j)t\right)\right) / \left(\frac{\hat{q}_j}{\hat{p}_j} \exp\left(-(\hat{p}_j + \hat{q}_j)t\right) + 1\right)$	(3)
Gompertz	$\hat{F}(j,t) = \exp\left(-\hat{a}_j \exp\left(-\hat{b}_j t\right)\right)$	(4)
Gamma/Shifted Gompertz	$\hat{F}(j,t) = \left(\left(1 - \exp(-\hat{b}_j t)\right)\right) / \left(\left(1 + \hat{a} \exp\left(-\hat{b}_j t\right)\right)^{\hat{c}_j}\right)$	(5)
Weibull	$\hat{F}(j,t) = 1 - \exp\left(-\left(t/\hat{a}_j\right)^{\hat{b}_j}\right)$	(6)
Non-parametric	$\hat{Y}(j,t) = (1/(2k+1)) \sum_{v=-k}^{k} Y_{j-1,t+v}$	(7)

4.2. Life cycle curves

For the adoption curve $F(\cdot)$ in Eq. (1), we use a set of parametric and non-parametric life cycle curves. We consider several alternative curve specifications to establish the usefulness of PROST beyond a single life cycle specification.

4.2.1. Parametric curve fitting

All parametric curves are listed in Table 3. With its steep decline, the adoption illustrated in Fig. 1 is typical for entertainment products, with various diffusion curves proposed in the literature (Clement et al., 2006). When forecasting with analogies, the Bass curve is the most frequently used (e.g. Bass et al., 2001; Goodwin et al., 2013; Lee et al., 2014). The cumulative adoption from the Bass curve for generation j is Eq. (3), where p_j and q_j are the coefficients of innovation and imitation for generation j, respectively (Bass, 1969). After each generation, we update the parameters p_j and q_j to account for any potential changes in buying behaviour over time, i.e., the rate of innovators may differ in the second generation due to the gaming experience in the previous generation. We consider two variants of the Gompertz curve, which provided the best fit on a mobile game diffusion study (Yi et al., 2019). The cumulative adoption is defined by Eq. (4), where a_j is the shape parameter and b_j the coefficient for scaling. The three parameters of the Gamma/Shifted Gompertz (G/SG) (Bemmaor, 1994) offer greater flexibility, and in the literature, there is evidence of good forecasting performance (Bemmaor & Lee, 2002; Bemmaor & Zheng, 2018; Meade & Islam, 2006). The shift is achieved by adding the c_j parameter in Eq. (5). Furthermore, we consider the Weibull curve that has been used for hedonic products (e.g. Moe & Fader, 2002). It is defined as in Eq. (6), where a_j is the shape parameter and b_j the parameter for scaling. Due to non-convergence, we abandoned the idea of using generational diffusion models that account for interaction between generations (e.g. Bass & Bass, 2001; Danaher et al., 2001; Norton & Bass, 1987).

All parametric life cycle models are estimated using non-linear least squares as available in the *diffusion* package (Schaer et al., 2024) for R (Team, 2024). Each curve is fitted only on data up until the release of the new game, including the required lead time, i.e., the parameters of a curve for generation j are estimated on generation $j-1$ up until period $T-l$. Since the sales history for some games spans several years, we truncated the right tail to facilitate the estimation of curve parameters, a common practice in life cycle models (e.g., Acimovic et al., 2019). Truncation occurs when the growth rate is less than 0.05 % of the weekly cumulative sales. On average, this limits the series to around 40 weeks.

4.2.2. Non-parametric curve fitting

Besides the parametric models, we introduce a non-parametric curve based on the centred moving average. In contrast to parametric curves, this approach can flexibly approximate any adoption shape. The curve is defined as in Eq. (7), where the order of the centred moving average is $2k+1$. We set k equal to 4 periods and interpolate the first and last $\frac{k}{2}$ periods of observed sales $Y_{j,t}$ from and to zero linearly.

4.3. Performance evaluation

Rather than using in-sample fit, which can lead to overfitting, we evaluate PROST models based on their predictive performance. Any

Table 4
Market potential estimation benchmarks.

Model	Equation	Required # coef.	Earliest estimation	Model description
B.1	$\hat{m}_j = m_{j-1} + \varepsilon_j$	0	$j = 2$	Naïve, previous generation market potential
B.2	$\hat{m}_j = c_0 \cdot m_{j-1}^{c_1} \cdot \varepsilon_j$	2	$j = 4$	Univariate equivalent to M.1
B.3	$\hat{m}_j = \alpha_0 + \alpha_1 m_{j-1} + \varepsilon_j$	2	$j = 4$	Univariate equivalent to M.1 (additive)

model that can predict has adequately approximated key elements of the underlying market structure. We generate forecasts for the market potential m_j, with $j \geq 2$, for each product, as described in Section 3. We then generate cumulative 52-week-ahead forecasts for the life cycle curves introduced in Section 4.2. In some cases, the evaluated forecast horizon is shorter, as there are no further observations in the dataset.

We consider several benchmarks to assess the gains from including PROST in the market potential estimation, presented in Table 4. A good way to assess predictive performance is to compare against simpler models that do not rely on parameter estimation (e.g., Armstrong, 2006). Such models are simple to implement, have low computational cost, and provide a good reference point for cross-study comparisons. In the case of pre-release forecasting, the Random Walk model corresponds to the observed market potential from the previous product generation (B.1). Furthermore, we include a univariate equivalent of M.1 (Table 2), which reduces to an autoregressive model on m_{j-1} (B.2). This directly displays the benefit added by PROST within the same model family. We also cover its additive autoregressive form (B.3). Note that the existing literature builds its baseline on product characteristics, with only two studies including past sales (Table 1). While it is possible to expand our models with additional product information, it would be impractical to do so. First, many of the product characteristics suggested in the literature, such as genre, age rating, or budget, are already represented through our analogy selection. Additionally, each extra variable increases the number of generations needed for estimation, making it a viable option only when the franchise has already matured substantially. This is especially true for categorical variables, which, in a regression context, result in $n-1$ dummy variables, likely necessitating further variable selection to limit model size; this can lead to unreliable results with limited data. For example, Schaer et al. (2022) found that at least 6 products were required to achieve reliable estimates in the context of competitive intelligence. We believe our model captures data characteristics similar to those found in existing literature. Because analogies naturally compress such information within past product generations, our approach is substantially more data-economical for companies with limited product histories. However, a direct comparison would likely favour models discussed in the literature, as they can leverage a larger data pool.

For each model, we assess the performance at each generation using the Geometric Mean Relative Absolute Error (GMRAE) (Armstrong & Collopy, 1992), due to its scale-independent properties, required for our dataset:

$$\mathrm{GMRAE}_{i,h} = \sqrt[n]{\prod_{i=1}^{n} \left(\frac{\mathrm{AE}_{i,r}}{\mathrm{AE}_{i,b}} \right)},$$

where $\mathrm{AE}_{i,h} = |\hat{y}_{i,t+h} - y_{i,t+h}|$ and r is the candidate model that gets divided by the Absolute Errors of the benchmark b. This is the Random Walk forecast (B.1) in Table 4, for each series i and horizon h. The error is summarised with the geometric mean across all n products. The geometric mean helps to make the scale more robust against outliers, which are common in video games. Moreover, GMRAE allows for an intuitive interpretation, where a GMRAE smaller than 1 implies that the evaluated model outperforms the selected benchmark by (1-GMRAE)*100 % (Ord et al., 2017). Furthermore, we track the bias in the predictions, providing insights into whether a model is prone to over- or under-forecasting. We report bias in the form of the Relative Median Error (RMdE):

$$\mathrm{RMdE}_{i,h} = Median_{i=1}^{n} \left(\frac{\mathrm{ME}_{i,r}}{|ME_{i,b}|} \right),$$

where the Mean Error of the candidate model r gets divided by the Mean Error of the Random Walk benchmark (B.1) b, and $\mathrm{ME} = \frac{1}{n} \sum_{h=1}^{n} (\hat{y}_{t+h} - y_{t+h})$. This is done to adjust for the scale of the errors. It follows that a candidate model with a positive value tends to over-forecast and vice versa, while normalisation by the benchmark allows for comparison across products. The metric is summarised using the median across all n products, as the geometric mean is unsuitable due to the possibility of negative values. With these two metrics, we ensure a comprehensive evaluation of both accuracy and systematic forecasting bias, which are crucial for practical implementation.

4.4. Results

4.4.1. Performance for the first week and end-of-life

While our model directly provides point forecasts for all 52 horizons (the complete life cycle), we focus on the first week (FW) and End-of-Life (EoL, cumulative sales at week 52) performance to address Hypothesis 1. In addition to testing our hypotheses, FW and EoL performance are important for several reasons. First, the first week represents the critical launch phase that most prior PRB literature targets (Table 1), allowing for a broad comparison with prior studies. Second, EoL reflects the overall market potential, as indicated by m_t in Eq. (1), which is essential for long-term decision-making. We examine performance across all horizons in Section 4.4.3.

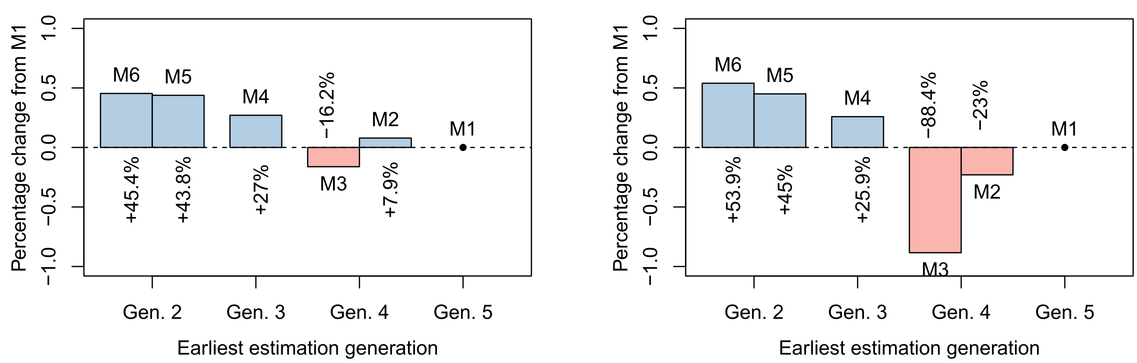
Table 5 shows FW and EoL performance across generations with a PROST lead time and a window size of 6 periods, which we found optimal across different lead times (discussed in Section 4.4.3). The number of estimated games for each generation is indicated in brackets for each column. Any empty cells reflect models that cannot be estimated due to insufficient observations. Models are identified as specified in Tables 2 and 4. The reported values are the combined performance for all life cycle models. Next to each error, an arrow indicates the average bias direction, pointing down or upwards for under-forecasting or over-forecasting, respectively. The best-performing model for each column is highlighted in bold.

The results reveal distinct performance patterns across model complexity levels, where the most restricted models (M.5 and M.6) consistently outperform their more complex counterparts. Specifically, M.6 outperforms all models for EoL and beats the best-performing benchmark, B.1, by 16.9 % (FW) and 21.3 % (EoL), although M.5 shows better FW performance for the third and older generations. Fig. 2 further demonstrates that parsimonious specifications can capture the underlying relationship between PROST and market potential when data is limited. The visualisation highlights the percentage differences in overall GMRAE between the restricted models and M.1, and organises the results by the first generation to which a model can be applied. M.5 and M.6 show substantial gains in performance and how early they can be used. M.4 shows moderate gains, while M.3, although restricted and usable before M.1, performs worse. M.2 offers minimal gains for the FW scenario and trails M.1 for EoL. In these two cases, the benefits of simpler estimation do not outweigh the restrictions in their functional forms. However, the sample size required to estimate M.1 limits its effectiveness. This effect is also evident in the benchmarks, where the

Table 5
Generation overall GMRAE performance across all adoption curves ($w = 6, l = 6$).

	Gen = 2 (57)		Gen = 3 (46)		Gen = 4 (28)		Gen = 5+ (63)		Overall (194)	
Model	FW	EoL	FW	EoL	FW	EoL	FW	EoL	FW	EoL
M.6	**0.848**↓	**0.872**↓	0.838↓	**0.678**↓	0.722↓	**0.725**↑	0.928↓	**0.896**↓	**0.831**↓	**0.787**↓
M.5	1.089↓	1.266↓	**0.813**↓	0.800↓	**0.678**↓	0.769↓	**0.889**↓	1.004↓	0.855↓	0.940↓
M.4			1.119↓	1.226↓	1.096↓	1.289↑	1.115↓	1.286↓	1.110↓	1.267↓
M.3					2.546↑	4.199↑	1.227↓	2.469↑	1.767↑	3.220↑
M.2					1.420↓	2.188↑	1.381↓	2.019↓	1.401↓	2.102↓
M.1							1.521↓	1.709↑	1.521↓	1.709↑
B.1	1.000↓	1.000↓	1.000↓	1.000↓	1.000↓	1.000↑	1.000↓	1.000↓	1.000↓	1.000↓
B.2					3.088↑	4.649↑	1.034↓	1.223↑	1.787↓	2.384↑
B.3					1.977↑	3.884↑	1.127↓	1.315↑	1.493↓	2.260↑

Brackets indicate the number of games within each generation; best performance per generation and horizon is in bold
FW = First Week, EoL = End-of-Life; ↑ = over-forecasting on average, ↓ = under-forecasting on average



(a) First week sales (b) End-of-life sales

Fig. 2. Percentage changes in overall GMRAE of restricted models from M.1. Models are organised by the first generation that can be applied to.

Table 6
GMRAE performance across all generations and forecasting horizons ($w = 6, l = 6$).

	Bass		Gompertz		Weibull		G/SG		CMA	
Model #	FW	EoL	FW	EoL	FW	EoL	FW	EoL	FW	EoL
M.6	0.846	0.778	**0.741**	0.790	0.742	**0.740**	0.877	1.062	0.774	0.772
M.5	0.924	0.957	**0.707**	0.949	0.738	0.963	0.949	1.122	0.800	**0.924**
M.1	1.562	1.740	1.074	1.431	1.401	1.926	1.684	2.858	**0.992**	**1.196**
B.1	1.000	1.000	0.904	1.016	0.952	0.988	1.047	1.307	**0.874**	**0.942**
B.2	1.684	1.737	1.239	1.870	3.736	5.651	1.483	2.673	**1.101**	**1.937**
B.3	1.179	1.676	1.189	1.721	2.602	5.193	1.531	2.810	**0.919**	**1.671**

All individual curves are assessed against the Bass Random Walk (B.1)

autoregressive-based B.2 and B.3 significantly improve for 5 or more generations without outperforming B.1.

Table 6 presents how individual life cycle curves with and without PROST compare to the Bass random walk (B.1). Overall, we can see that the *CMA*, *Gompertz*, and *Weibull* curves perform better. In contrast, the *G/SG* provides the weakest forecasting performance, especially for the EoL scenario.

To further confirm the suggested finding that PROST adds predictive value, we test whether differences are significant and not due to randomness. As the GMRAE errors are not normally distributed, we apply non-parametric Friedman and post-hoc Nemenyi tests (see Hollander et al., 2014). We use them to determine whether any errors from models M.6 and B.1 are significantly different, utilising the *tsutils* package for R (Kourentzes, 2023). We find that, in both forecast scenarios, all models differ significantly. Therefore, in relation to Hypothesis 1, we find that PROST contains a predictive value not only for the initial week but also for the entire market potential.

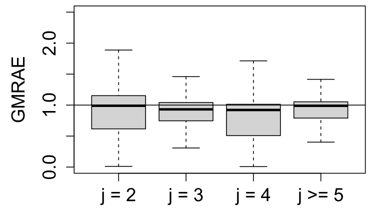
In Hypothesis 3, we hypothesise whether the predictive content of PRB changes across product generations. Since the sample sizes differ across generations, the Friedman test is inappropriate; we replace it with a nonparametric Kruskal-Wallis test. We find that the errors of M.6 differ significantly across generations only in the EoL case. In Fig. 3, the box plots show the error spreads of M.6. However, there is no clear trend in the distribution of errors that would suggest PROST becomes more informative as a franchise matures.

4.4.2. Bias evaluation of PROST

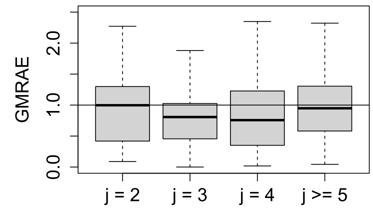
The observed bias, indicated by arrows in Table 5, depicts consistent under-forecasting by the best-performing models, M.5 and M.6. Table 7 quantifies this bias across generations for M.6, M.5, and B.1 and highlights the difference between these models (bottom 2 rows). All models are under-forecast, less frequently for EoL than FW sales, and without any clear trend as franchises mature.

4.4.3. Leading properties and impact of window size

In Hypothesis 2, we raise the question about the leading property of PROST. Thus far, all presented results correspond to a window size of 6 weeks and a lead time of 6 weeks. We find that different window sizes



(a) First week sales



(b) End-of-life sales

Fig. 3. Distribution of GMRAE errors across generations for M.6.

Table 7
Percentage of under-forecasting across the different generations ($w = 6, l = 6$).

	Gen.2		Gen. 3		Gen. 4		Gen. 5 +	
Model #	FW	EoL	FW	EoL	FW	EoL	FW	EoL
M.6	76 %	71 %	70 %	61 %	70 %	49 %	76 %	59 %
M.5	73 %	67 %	71 %	61 %	76 %	52 %	79 %	60 %
B.1	76 %	78 %	65 %	64 %	66 %	46 %	71 %	54 %
M.6 - B.1	0 %	−7 %	5 %	−3 %	4 %	3 %	5 %	5 %
M.5 - B.1	−3 %	−11 %	6 %	−3 %	10 %	6 %	8 %	6 %

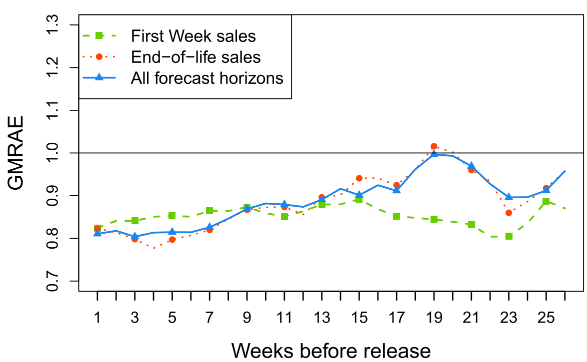


Fig. 4. Forecast performance of M.6 for all scenarios at different lead times ($w = 6$).

(w) reduce forecast accuracy across all lead times, suggesting that window size is application-dependent and should be determined through out-of-sample testing in practice.

In Fig. 4, the dashed green line and dotted red line illustrate the performance of M.6 across different lead times (l) for FW and EoL, respectively. In the FW scenario, PROST has a predictive value 26 weeks before product release. EoL is estimated to be better than FW predictions for the 9 weeks prior to release. FW observes one of the best performances 23 weeks in advance, while EoL forecasts have the lowest errors around 3-5 weeks prior to release. Note that the sample size is reduced to 170 games for the 26-week lead-time case, as some games lacked a search traffic signal.

So far, we have only considered the performance of the first and last weeks of sales. However, decision-makers might also require forecasts over the entire life cycle evolution, for instance, to support the planning of advertising campaigns. The forecast scenario across all horizons (solid blue line in Fig. 4) closely resembles the EoL performance. In both the EoL and all-horizon cases, the performance steadily improves towards release 19 weeks prior to the release date. As a further reference, the overall GMRAE performance at $l = 6$ is 0.814 and 0.910 for M.6 and M.5, respectively.

Finally, we note that the literature has also suggested approximating PROST using splines (Xiong & Bharadwaj, 2014) or growth curves (Kulkarni et al., 2012) until the product launch, e.g., forecasting PROST 6 steps ahead when $l = 6$. We investigated a similar approach by predicting PROST using exponential smoothing models, but found no significant benefit.

4.4.4. Comparison with ad spend

We compared PROST against Pre-Release Ad Spend (PRAS), given publishers' substantial pre-release advertising investments (Marchand & Hennig-Thurau, 2013), and PROST's partial capture of these expenditures (Hu et al., 2014; Srinivasan et al., 2016). This poses the question of whether PRAS, on its own, would provide similar insights. To test this assumption, we obtained PRAS information from the Kantar Media Ad$pender database (similar to, Kim & Hanssens, 2017; Xiong & Bharadwaj, 2014). We were unable to obtain ad spending data for four games, so the following comparison is based on 191 games.

As with PROST, the initial intention was to assess PRAS's predictive information across different windows and lead times. However, many games only observe a few weeks of pre-release ad spend, making a window-based approach unsuitable. Therefore, the results in Table 8 use the entire pre-release ad spend available (same as Xiong & Bharadwaj, 2014). This also considers that advertisers have prior expectations of market potential, implying some predictive ability.

Table 8 shows the percentage change in forecasting performance for models M.6 to M.1 using PRAS in direct comparison to its PROST equivalents. Almost all PRAS models exhibit substantial losses in predictive performance (indicated by negative values), with only a few showing marginal improvements. However, these improvements do not outperform the best PROST model and are close to, or even worse than, the random walk model (B.1). Consequently, we abandoned the idea of investigating models that simultaneously use PROST and PRAS information. Furthermore, the additional cost in degrees of freedom would make such models even more cumbersome, given the limited data points available for estimation.

5. Discussion and managerial consequences

This study has explored the predictive power of PRB in conjunction with analogies for marketing and operations decisions by empirically testing three theoretically and practically relevant hypotheses. We accept Hypothesis 1, as PROST predicts both early sales and overall market potential, extending traditional first-week forecasting methods. Regarding the lead times examined in Hypothesis 2, our results show that PROST performance gains are available for all scenarios 18 weeks before release. The lead time benefit of knowing first-week sales is even more substantial, with significant improvements 26 weeks in advance. Finally, when we examined generational differences in performance, we partially rejected Hypothesis 3 as we only found a significant variation in PROST's predictive ability for the EoL case. In addition, we were unable to visually identify any systematic trend, indicating that PROST does not become more or less informative as a franchise matures. This may explain why the literature has treated sequels as a single explanatory variable rather than individual generations.

PROST carries different information for the first week and end-of-life sales. It may seem counterintuitive that long-term estimates consistently outperform first-week sales when forecasting with PROST within

Table 8
GMRAE improvement for PRAS across all adoption curves against PROST ($w = 6, l = 6$).

	Gen. = 2 (57)		Gen = 3 (45)		Gen = 4 (28)		Gen = 5+ (62)		Overall (191)	
Model #	FW	EoL	FW	EoL	FW	EoL	FW	EoL	FW	EoL
M.6	−95 %	−133 %	−103 %	−124 %	−14 %	−22 %	−74 %	−145 %	−65 %	−97 %
M.5	−183 %	−252 %	−219 %	−284 %	−50 %	−46 %	−194 %	−398 %	−147 %	−209 %
M.4			−31 %	−19 %	−5 %	**1 %**	−20 %	−44 %	−18 %	−20 %
M.3					54 %	62 %	−22 %	−19 %	**6 %**	**14 %**
M.2					**2 %**	0 %	−22 %	−45 %	−10 %	−22 %
M.1							**11 %**	−31 %	**11 %**	−31 %

Brackets indicate the number of games within each generation bucket; Models outperforming equivalent PROST model in bold

9 weeks of release. The literature suggests that pre-release advertising (PRAS) and publicity have a positive impact on long-term sales (e.g. Burmester et al., 2015; Elberse & Anand, 2007). We show that this relationship also holds for PROST. To some extent, this is expected, given the close relationship between marketing efforts and PRB (Kim & Hanssens, 2017), but our findings demonstrate that PROST's predictive value extends beyond the first week, a fact the literature has overlooked.

Forecast improvements are achieved using search traffic prior to the product's release date. Notably, the first week of sales forecasts, including PROST, outperforms our benchmarks for all lead times up to 26 weeks. However, for the entire market potential, we find that performance improves continuously from week 17 onward. We argue that the closer to release, the greater the impact on long-term sales, as consumers' certainty increases. Moreover, video game advertising campaigns typically begin around 9 weeks prior to release (Marchand et al., 2017). They may explain why PROST becomes more informative as release approaches, reflecting consumer response to advertising spend. However, the feasible lead time depends on the product type, as consumers invest varying amounts of effort in researching different products. For example, consumers spend more time on high-involvement purchases, such as cars and video games, than on everyday items (Bhatnagar & Ghose, 2004).

Significant differences across generations were observed only in end-of-life and all-horizon scenarios. A wider error spread between long-term scenarios is expected as additional factors become relevant, such as word of mouth or post-release advertising, which affect overall adoption. The literature argues that the uncertainty of product adoption is reduced over the lifespan of a franchise and directly cannibalises advertising effects (Hennig-Thurau & Houston, 2019). This would intuitively suggest that PRB has become a less reliable indicator over time, as consumers are increasingly familiar with the franchise, possibly leading to reduced search activity and misleadingly indicating lower sales. However, our findings do not support this interpretation. One possible reason is that internet searches become more focused over time: while consumers familiar with the product might still check specific details, such as release dates or prices, there may be fewer broad exploratory searches from consumers who ultimately do not make a purchase.

Consistent with previous studies, we find that ad spend has limited predictive value (Foutz & Jank, 2010; Kulkarni et al., 2012; Xiong & Bharadwaj, 2014). Our ad-based models performed similarly to a random walk, indicating that model combinations that include advertising data are unlikely to be beneficial (see also, Foutz & Jank, 2010, where combinations perform worse than PRB on its own). This limited effectiveness can be partially explained by research showing that an increase in ad spend leads to only a small increase in search traffic volume (Kim & Hanssens, 2017) and has a minimal overall impact on PRB (Schreiner et al., 2025). Therefore, we conclude that it is vital for companies to monitor and seed the PRB and its various sources, as this better reflects the underlying interest in new products.

From a modelling perspective, we note that the most parsimonious alternatives (M.5 & M.6) have the best predictive performance of new product sales. While model M.5 has a PROST coefficient set to 1, setting it to 0.5 in M.6 greatly reduces error variance. Because a multiplicative formulation is prone to generating larger outliers due to sensitivity to shifts in PROST, the damping M.6 largely overcomes this issue. This shows that (i) parsimonious PROST models are superior to more complex counterparts and (ii) further restriction of coefficients, as motivated by shrinkage in M.6, increases predictive performance even further.

Table 9
Market potential estimation for first-generation.

Model	Equation
MF.6	$\hat{m}_{j=1} = \frac{1}{F}\sum_{f=1}^{F}((\mathrm{PROST}_{j=1} - \mathrm{PROST}_{f,j=2})^{\frac{1}{2}} \cdot m_{f,j=2})\varepsilon_j$
BF.1	$\hat{m}_{j=1} = \frac{1}{F}\sum_{f=1}^{F} m_{f,j=1} + \varepsilon_j$
BF.2	$\hat{m}_{j=1} = \frac{1}{F}\sum_{f=1}^{F}((\mathrm{PROST}_{f,j=1} - \mathrm{PROST}_{f,j=2})^{\frac{1}{2}} \cdot m_{f,j=2})\varepsilon_j$

Table 10
GMRAE performance for first-generation across all adoption curves ($w = 6, l = 6$).

	MF.6	BF.1	BF.2
FW	**0.920**	1.000	1.322
EoL	**0.924**	1.000	1.085

5.1. *Forecasting the first generation*

Our argument for using sequential rather than longitudinal data is that it simplifies the process of identifying analogies. However, a significant limitation is the requirement of at least one life cycle before any estimation can be done. In the following, we elaborate on how to extend our proposed approach to predict first-generation games.

We build on the best-performing M.6 model. First, we calculate the difference between the PROST for the target first-generation product and the second-generation PROST of already existing products of the same publisher, $\mathrm{PROST}_{j=1} - \mathrm{PROST}_{f,j=2}$, where $f = 1, \ldots, F$ and F is the number of products from the same publisher that has a second generation. We then average across all estimates from the F products to obtain $\hat{m}_{j=1}$ (MF.6 Table 9). The alternative, first averaging and then estimating, as suggested by Hu et al. (2019), leads to worse performance.

We introduce two new benchmarks that utilise publisher-wide data in Table 9. BF.1 is the average market potential for the first generation across all available franchises from the same publisher. BF.2 follows the idea of MF.6 but strictly uses the publisher PROST data for both the first and second generations without using the PROST for the first generation of the target product. BF.1 is a substitute for our benchmark B.1 introduced in Table 4, which is unsuitable for first-generation forecasting. BF.2 indicates whether historical PROST can improve estimation over publisher-based information. Note that BF.1 and BF.2 are not well suited to assess the performance for later instalments, as the number of mature games from the same publisher diminishes rapidly.

We estimate $\hat{m}_{j=1}$ using the same dataset. We had no additional franchises from the same publisher available for two instances. To maintain the overall sample size, we each randomly selected four franchises from different publishers. Table 10 reports the GMRAE performance across all

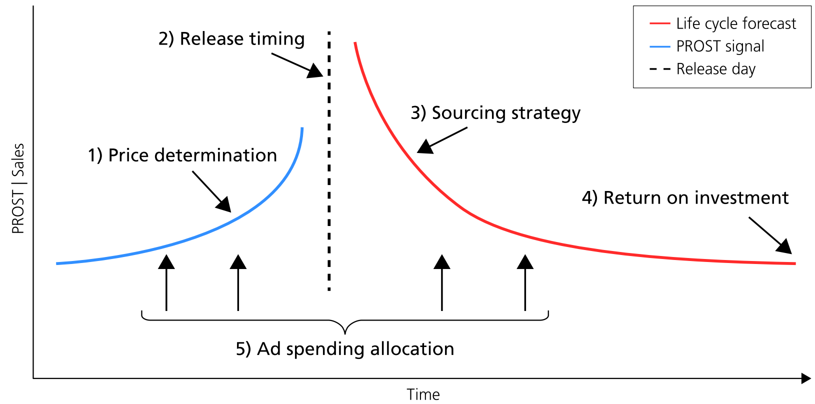


Fig. 5. Decision points along the video game life cycle.

adoption curves. We observe that MF.6 outperforms the benchmarks for the FW and EoL scenarios, providing evidence that even without searching for optimal analogies, PROST yields improvements over using data from other franchises alone.

5.2. Impact for practice

Data-driven analytics has become an integral part of the video game industry. Developers and publishers benefit from leveraging data to make more informed decisions. Fig. 5 highlights some of the critical decision points within a video game life cycle where our model adds value:

1. Price determination. Companies can opt for a penetration or price-skimming strategy that will impact the overall product adoption (see Spann et al., 2015). Video games are often launched at suboptimal prices, particularly for smaller franchises. This severely harms their revenue potential and reduces their ability to offer discounts in the later life-cycle stage. For example, seasonal sales on Steam, the largest online distribution platform, are a key revenue driver after launch (Dealessandri, 2020). Forecasted adoption rate and market potential are required to initialise model parameters for price optimisation (e.g., Zhang et al., 2022). Similarly, product bundles are an attractive way for publishers to increase profits through indirect price discrimination (see Hennig-Thurau & Houston, 2019). Forecasting lifetime sales can support a holistic decision-making process where pricing, discounts, and bundling can be considered across the product's life cycle to maximise returns.

2. Return on investment. Accurate life-cycle forecasts inform investment decisions, with developers like Remedy using lifetime sales targets to determine project continuation (Yin-Poole, 2024). Take-Two recently sold one of its indie game divisions because it did not meet expected sales targets, despite profitability (Chalk, 2024). Knowing that PROST is a helpful tool to predict life cycle sales allows managers to i) start countermeasures earlier and ii) adjust investments into future developing activities of the series. The accuracy of return-on-investment forecasts will increase post-launch; however, the time required to collect sufficient data limits the remedies management can implement.

3. Sourcing strategy. For short-life cycle products, dual-sourcing is a common strategy to increase production flexibility. In other words, the company has both a low-cost, long-lead supplier and a local supplier for short-lead-time replenishments (e.g., Calvo & Martínez-de Albéniz, 2016). A similar concept exists for hosting video games, where hybrid solutions extend cost-efficient bare-metal servers with virtual machines in the cloud to handle demand spikes (see Pandey, 2024, for a use case of how hybrid server solutions operate). Our model's improved forecast performance and long-lead-time capability enable us to shift demand towards the cheaper supplier option.

4. Release timing. Research suggests that managing the release timing can be more important than the overall price for short-life cycle products (Engelstätter & Ward, 2018), especially in times when even high-level video games struggle to increase prices (Evans-Thirlwell, 2023). For pricing and release timing decisions, it is vital to understand the competitive standing against competitors' expected success. Schaer et al. (2022), show that PROST can be used as a proxy to estimate competitors' market potential.

5. Ad spending allocation. For video game publishers, allocation of ad spending is crucial, and our research suggests that managing PROST is important. The study by Marchand and Hennig-Thurau (2013) suggests that publishers invested more in post-release campaigns during the time frame covered by our data. This finding is counterintuitive to typical patterns of product adoption and conflicts with existing research, which has shown that pre-release video game advertising is marginally more effective than post-release advertising (Hennig-Thurau & Houston, 2019). Several studies recommend a greater focus on pre-release campaigns (Burmester et al., 2015; Kim & Hanssens, 2017; Mandler et al., 2025). Additionally, the timing and design of pre-release advertisement significantly influence the buzz generated around a game (Nguyen & Chaudhuri, 2019; Schroll & Grohs, 2019). In particular, the findings of Schreiner et al. (2025) suggest that mass media advertising can create buzz only under certain specific circumstances, while social media posts tend to be much more effective.

5.3. Limitations and further areas of research

While we focus on PROST, comprehensive PRB analysis could incorporate additional signals, such as anticipatory communication and experiential activities (Houston et al., 2018). For example, Schreiner et al. (2025) suggest that for movies, pre-release communication drives sales the most, with search and participation playing a supporting role. At the same time, the surveyed studies in Section 2 are somewhat inconclusive. Therefore, we cannot rule out the possibility that other sources may contain more predictive information. Specifically, pre-orders (Moe & Fader, 2002; Mostard et al., 2011) and unstructured PRB data, such as images (Skenderi et al., 2024), dynamic textual information (Zhong & Schweidel, 2020), and sentiment analysis (Cui et al., 2018), offer promising avenues for future exploration within our proposed approach. While the digital landscape has expanded the availability of consumer engagement data, significant challenges remain in practice. For example, in the video game ecosystem, consumer engagement occurs across multiple touchpoints, from in-game behaviour and marketing responses to gaming community interactions. However, integrating these different data sources presents substantial methodological and practical challenges (Welden et al., 2025). Additionally, many of these data sources are not widely accessible, limiting their broader application. Nonetheless, we believe that our parsimonious method provides a solid foundation for investigating additional data sources and assessing their effectiveness in forecasting the adoption of sequential and new products.

However, even when focusing on publicly available data sources, such as Google Trends, keyword identification can be challenging (see Schaer et al., 2019, for a discussion). Despite this challenge, we believe PROST has broader applications beyond gaming and entertainment. For example, in inventory management, inaccurate predictions in the launch phase notoriously lead to stock-outs or higher obsolescence costs (Babai et al., 2025). In particular, we see potential in exploring the application for manufacturers of popular electronic goods, such as digital cameras, mobile phones, and computer components, as well as in the automotive and luxury fashion industries, where consumers actively seek information about upcoming releases. In contrast, utilitarian or niche products may observe limited measurable product anticipation (Houston et al., 2018; Mandler et al., 2025). For those products, crowd-sourced keywords, as suggested by Brynjolfsson et al. (2016), are a possible solution worth exploring further.

By definition, PRB leads any product launch, which ensures exogeneity. A valid concern remains the potential influence of ad spend, where the expected market potential influences advertising decisions. However, in line with the literature, we find little evidence that advertising expenditure is a strong indicator of sales outcomes. Moreover, our model's tendency to under-forecast demand could be complemented by expert judgment, which tends to over-forecast new product demand, according to a survey of practitioners (Kahn & Chase, 2018). Although the literature on combining statistical pre-launch forecasting models and expert judgment is sparse (Perera et al., 2019), such an approach may improve forecast accuracy by mitigating opposing biases (Fildes et al., 2025). There is also room to investigate the extent to which PROST can support judgmental decision-making, as evidence from the literature suggests that contextual information can lead to better forecasts (Seifert et al., 2015).

6. Conclusions

Better forecasts can support better operations and planning, particularly for products with short life cycles, such as video games. Our proposed approach can facilitate algorithmic development to optimize operations, which are often done ad hoc in practice. Our research addresses this issue by introducing a new methodology for forecasting product life cycles, leveraging PROST to provide concrete decision support for pricing, marketing allocation, and launch timing. We contribute to new product forecasting literature in two ways:

First, by establishing PROST's relationship with the pre-launch market potential. In the literature on analogy-based forecasting, market potential is often determined by expert judgment and is known to be prone to bias. Our suggested methodology offers an alternative approach to determining a new product's market potential in a more data-driven manner. PROST offers freely available, high-frequency data for tracking shifts in consumer interest and competitive standing, enabling timely forecast updates (Meeran et al., 2017). The lead time during which such predictive PROST information is available allows decision-makers to adjust their marketing and operational planning continuously until release.

Second, our proposed methodology is very data-economical, requiring only a single franchise. In contrast, the majority of PRB studies rely on large datasets spanning many products and companies (e.g., Lehrer & Xie, 2021; Xiong & Bharadwaj, 2014). Companies are unlikely to have a large pool of previously introduced products. Instead, our approach utilises the natural analogy for new products that are part of a brand extension and incremental product innovations. The proposed M.6 model requires only two PROST generations to gain valuable insights into market potential. Machine learning approaches would require a substantially larger sample, which would make them challenging for future work. We provide routes forward for expanding this approach to cases without previous product generations and evaluate the usefulness of Ad Spend information when PROST is available. In both cases, PROST is found to be beneficial.

Finally, the proposed models provide full life cycle predictions. This contrasts with the PRB literature, which focused on specific periods after launch. The proposed models provide forecasts throughout the life cycle with a single estimation, while existing approaches must be recalibrated and re-estimated for each lead time. This offers a practical tool for sales planning and creates opportunities for practice and research, specifically in providing the necessary inputs towards developing prescriptive models, which are under-represented in the literature relative to the size of the video games industry.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors used Claude Sonnet to assist with spelling and syntax. The authors thoroughly reviewed all suggestions, taking full responsibility for the publication's content.

Acknowledgement

The authors would like to thank the Editor and the two anonymous reviewers for their insightful and detailed feedback, which greatly contributed to improving this manuscript. We are also grateful to Christophe Van den Bulte and an anonymous AE at Marketing Science for their advice on the positioning and framing of our work. Finally, we thank Florian Dost and Nigel Meade for their helpful early-stage comments.

CRediT authorship contribution statement

Oliver Schaer: Writing – review & editing, Writing – original draft, Visualization, Methodology, Data curation, Formal analysis, Conceptualization; **Nikolaos Kourentzes:** Writing – review & editing, Visualization, Supervision, Methodology, Formal analysis, Conceptualization; **Robert Fildes:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Acimovic, J., Erize, F., Hu, K., Thomas, D. J., & Van Mieghem, J. A. (2019). Product life cycle data set: Raw and cleaned data of weekly orders for personal computers. *Manufacturing & Service Operations Management*, *21* (1), 171–176.

Armstrong, J. S. (2006). Findings from evidence-based forecasting: Methods for reducing forecast error. *International Journal of Forecasting*, *22* (3), 583–598.

Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, *8* (1), 69–80.

Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. In *Web intelligence and intelligent agent technology (WI-IAT), 2010 IEEE/WIC/ACM International conference on* (pp. 492–499). (*vol. 1*).

Baardman, L., Levin, I., Perakis, G., & Singhvi, D. (2018). Leveraging comparables for new product sales forecasting. *Production and Operations Management*, *27* (12), 2340–2343.

Babai, M. Z., Syntetos, A. A., & Teunter, R. H. (2025). Fifty years of inventory research from a forecasting perspective. *European Journal of Operational Research*. pre-print, https://doi.org/10.1016/j.ejor.2025.07.003.

Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, *15* (5), 215–227.

Bass, F. M., Gordon, K., Ferguson, T. L., & Githens, M. L. (2001). DirecTV: Forecasting diffusion of a new technology prior to product launch. *Interfaces*, *31* (3), S82–S93.

Bass, P. I., & Bass, F. M. (2001). Diffusion of technology generations: A model of adoption and repeat sales. *Working Paper*, (pp. 1–29). https://api.semanticscholar.org/CorpusID:51987514.

Belvedere, V., & Goodwin, P. (2017). The influence of product involvement and emotion on short-term product demand forecasting. *International Journal of Forecasting*, *33* (3), 652–661.

Bemmaor, A. C. (1994). Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity. In G. Laurent, G. L. Lilien, & B. Pras (Eds.), *Research traditions in marketing* (pp. 201–229). Dordrecht: Springer.

Bemmaor, A. C., & Lee, J. (2002). The impact of heterogeneity and ill-conditioning on diffusion model parameter estimates. *Marketing Science*, *21* (2), 209–220.

Bemmaor, A. C., & Zheng, L. (2018). The diffusion of mobile social networking: Further study. *International Journal of Forecasting*, *34* (4), 612–621.
Bhatnagar, A., & Ghose, S. (2004). Online information search termination patterns across product categories and consumer demographics. *Journal of Retailing*, *80* (3), 221–228.
Brynjolfsson, E., Geva, T., & Reichman, S. (2016). Crowd-squared: Amplifying the predictive power of search trend data. *MIS Quarterly*, *40* (4), 941–962.
Burmester, A. B., Becker, J. U., van Heerde, H. J., & Clement, M. (2015). The impact of pre- and post-launch publicity and advertising on new product sales. *International Journal of Research in Marketing*, *32* (4), 408–417.
Calvo, E., & Martínez-de Albéniz, V. (2016). Sourcing strategies and supplier incentives for short-life-cycle goods. *Management Science*, *62* (2), 436–455.
Chalk, A. (2024). Take-two has sold private division to an unnamed buyer: 'We are top-ten hit makers around here,' and private division wasn't making them. PCGamer.com, accessed on 28/11/2024. https://tinyurl.com/33x2fcb3.
Chintagunta, P. K., & Lee, J. (2012). A pre-diffusion growth model of intentions and purchase. *Journal of the Academy of Marketing Science*, *40* (1), 137–154.
Chumnumpan, P., & Shi, X. (2019). Understanding new products' market performance using google trends. *Australasian Marketing Journal (AMJ)*, *27* (2), 91–103.
Clement, M., Fabel, S., & Schmidt-Stolting, C. (2006). Diffusion of hedonic goods: A literature review. *International Journal on Media Management*, *8* (4), 155–163.
Craig, C. S., Greene, W. H., & Versaci, A. (2015). E-Word of mouth: Early predictor of audience engagement. *Journal of Advertising Research*, *55* (1), 62–72.
Cui, R., Gallino, S., Moreno, A., & Zhang, D. J. (2018). The operational value of social media information. *Production and Operations Management*, *27* (10), 1749–1769.
Danaher, P. J., Hardie, B. G. S., & Putsis, W. P., Jr (2001). Marketing-mix variables and the diffusion of successive generations of a technological innovation. *Journal of Marketing Research*, *38* (4), 501–514.
Dealessandri, M. (2020). How to price your game and thrive during steam sales. GamesIndustry.biz, accessed on 28/11/2024. https://tinyurl.com/mr297xv5.
Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21* (4), 23–45.
Ding, C., Cheng, H. K., Duan, Y., & Jin, Y. (2017). The power of the "like" button: The impact of social media on box office. *Decision Support Systems*, *94*, 77–84.
Divakaran, P. K. P., Palmer, A., Søndergaard, H. A., & Matkovskyy, R. (2017). Pre-launch prediction of market performance for short lifecycle products using online community data. *Journal of Interactive Marketing*, *38*, 12–28.
Du, R. Y., & Kamakura, W. A. (2012). Quantitative trendspotting. *Journal of Marketing Research*, *49* (4), 514–536.
Elberse, A., & Anand, B. (2007). The effectiveness of pre-release advertising for motion pictures: An empirical investigation using a simulated market. *Information Economics and Policy*, *19* (3), 319–343.
Eliashberg, J., Jonker, J.-J., Sawhney, M. S., & Wierenga, B. (2000). MOVIEMOD: An implementable decision-support system for prerelease market evaluation of motion pictures. *Marketing Science*, *19* (3), 226–243.
Engelstätter, B., & Ward, M. R. (2018). Strategic timing of entry: Evidence from video games. *Journal of Cultural Economics*, *42* (1), 1–22.
ESA (2023). 2023 essential facts about the U.S. video game industry. http://tinyurl.com/2nwf357f.
Evans-Thirlwell, E. (2023). GTA 6 publisher ceo thinks videogame prices are "very, very low" for what they offer. Rockpapershotgun.com, accessed on 28/11/2024 https://tinyurl.com/3h949ezc.
Feiler, D., & Tong, J. (2022). From noise to bias: Overconfidence in new product forecasting. *Management Science*, *68* (6), 4685–4702.
Fildes, R., Goodwin, P., & De Baets, S. (2025). Forecast value added in demand planning. *International Journal of Forecasting*, *41*, 649–669.
Foutz, N. Z., & Jank, W. (2010). Research note– prerelease demand forecasting for motion pictures using functional shape analysis of virtual stock markets. *Marketing Science*, *29* (3), 568–579.
Fu, Y., & Fisher, M. (2023). The value of social media data in fashion forecasting. *Manufacturing & Service Operations Management*, *25* (3), 1136–1154.
Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., & Watts, D. J. (2010). Predicting consumer behavior with web search. *Proceedings of the National Academy of Sciences*, *107* (41), 17486–17490.
Goodwin, P., Dyussekeneva, K., & Meeran, S. (2013). The use of analogies in forecasting the annual sales of new electronics products. *IMA Journal of Management Mathematics*, *24* (4), 407.
Goodwin, P., Meeran, S., & Dyussekeneva, K. (2014). The challenges of pre-launch forecasting of adoption time series for new durable products. *International Journal of Forecasting*, *30* (4), 1082–1097.
Green, K. C., & Armstrong, J. S. (2015). Simple versus complex forecasting: The evidence. *Journal of Business Research*, *68* (8), 1678–1685. Special Issue on Simple Versus Complex Forecasting.
Hastie, T., Tibshirani, R., & Wainwright, M. (2015). Statistical learning with sparsity. The lasso and generalizations. Monographs on Statistics and Applied Probability 143. Boca Raton: CRC Press.
Hennig-Thurau, T., & Houston, M. B. (2019). Entertainment science. Data analytics and practical theory for movies, games, books, and music. Switzerland: Springer.
Hennig-Thurau, T., Houston, M. B., & Heitjans, T. (2009). Conceptualizing and measuring the monetary value of brand extensions: The case of motion pictures. *Journal of Marketing*, *73* (6), 167–183.
Hollander, M., Wolfe, D. A., & Chicken, E. (2014). Nonparametric statistical methods. (3rd ed.). New Jersey: Wiley.
Houston, M. B., Kupfer, A.-K., Hennig-Thurau, T., & Spann, M. (2018). Pre-release consumer buzz. *Journal of the Academy of Marketing Science*, *46* (2), 338–360.
Hu, K., Acimovic, J., Erize, F., Thomas, D. J., & Van Mieghem, J. A. (2019). Forecasting new product life cycle curves: Practical approach and empirical analysis. *Manufacturing & Service Operations Management*, *21* (1), 66–85.
Hu, Y., Du, R. Y., & Damangir, S. (2014). Decomposing the impact of advertising: Augmenting sales with online search data. *Journal of Marketing Research*, *51* (3), 300–319.
Kahn, K. B., & Chase, C. W. (2018). The state of new-product forecasting. *Foresight: The International Journal of Applied Forecasting*, *3* (51), 24–31.
Kim, H., & Hanssens, D. M. (2017). Advertising and word-of-mouth effects on pre-launch consumer interest and initial sales of experience products. *Journal of Interactive Marketing*, *37*, 57–74.
Kim, T., Hong, J., & Kang, P. (2015). Box office forecasting using machine learning algorithms based on SNS data. *International Journal of Forecasting*, *31* (2), 364–390.
Kim, T., Hong, J., & Koo, H. (2013). Forecasting diffusion of innovative technology at pre-launch: A survey-based method. *Industrial Management & Data Systems*, *113* (6), 800–816.
Kourentzes, N. (2023). tsutils: Time series exploration, modelling and forecasting. R package version 0.9.4. https://cran.r-project.org/package=tsutils.
Kulkarni, G., Kannan, P. K., & Moe, W. (2012). Using online search data to forecast new product sales. *Decision Support Systems*, *52* (3), 604–611.
Lee, H., Kim, S. G., Park, H.-w., & Kang, P. (2014). Pre-launch new product demand forecasting using the bass model: A statistical and machine learning-based approach. *Technological Forecasting and Social Change*, *86*, 49–64.
Lee, J., Cho, Y., Lee, J.-D., & Lee, C.-Y. (2006). Forecasting future demand for large-screen television sets using conjoint analysis with diffusion model. *Technological Forecasting and Social Change*, *73* (4), 362–376.
Lehrer, S. F., & Xie, T. (2021). The bigger picture: Combining econometrics with analytics improves forecasts of movie success. *Management Science*, *68* (1), 189–210.
Lei, D., Hu, H., Geng, D., Zhang, J., Qi, Y., Liu, S., & Shen, Z.-J. M. (2023). New product life cycle curve modeling and forecasting with product attributes and promotion: A Bayesian functional approach. *Production and Operations Management*, *32* (2), 655–673.
Lenk, P. J., & Rao, A. G. (1990). New models from old: Forecasting product adoption by hierarchical Bayes procedures. *Marketing Science*, *9* (1), 42–53.
Loveridge, S. (2025). GTA 6: Everything we know so far. Gamesradar.com, accessed on 24/01/2025. https://tinyurl.com/5as24d5w.
Machkovech, S. (2024). Why animal well's home-brewed engine was key to its success. Gamedeveloper.com, accessed on 24/01/2025 https://tinyurl.com/22ufw3fc.
Mahajan, V., Sharma, S., & Sharma, S. (1986). A simple algebraic estimation procedure for innovation diffusion models of new product acceptance. *Technological Forecasting and Social Change*, *30* (4), 331–345.
Mandler, T., Kupfer, A.-K., Hennig-Thurau, T., Schauerte, R., & Cziehso, G. P. (2025). The contagious nature of pre-release consumer buzz: How observing other consumers' anticipation for a new product influences adoption. *Journal of Product Innovation Management*, *42*, 1044–1067.
Marchand, A. (2016). The power of an installed base to combat lifecycle decline: The case of video games. *International Journal of Research in Marketing*, *33* (1), 140–154.
Marchand, A., & Hennig-Thurau, T. (2013). Value creation in the video game industry: Industry economics, consumer benefits, and research opportunities. *Journal of Interactive Marketing*, *27* (3), 141–157.
Marchand, A., Hennig-Thurau, T., & Wiertz, C. (2017). Not all digital word of mouth is created equal: Understanding the respective impact of consumer reviews and microblogs on new product success. *International Journal of Research in Marketing*, *34* (2), 336–354.
Markham, S. K., & Lee, H. (2013). Product development and management association's 2012 comparative performance assessment study. *Journal of Product Innovation Management*, *30* (3), 408–429.
Markovitch, D. G., Steckel, J. H., Michaut, A., Philip, D., & Tracy, W. M. (2014). Behavioral reasons for new product failure: Does overconfidence induce overforecasts? *Journal of Product Innovation Management*, *32* (5), 825–841.
Meade, N., & Islam, T. (2006). Modelling and forecasting the diffusion of innovation - a 25-year review. *International Journal of Forecasting*, *22* (3), 519–545.
Meeran, S., Jahanbin, S., Goodwin, P., & Neto, J. Q. F. (2017). When do changes in consumer preferences make forecasts from choice-based conjoint models unreliable? *European Journal of Operational Research*, *258* (2), 512–524.
Moe, W. W., & Fader, P. S. (2002). Fast-track: Article using advance purchase orders to forecast new product sales. *Marketing Science*, *21* (3), 347–364.
Moon, S., Kim, J., Bayus, B. L., & Yi, Y. (2016). Consumers' pre-launch awareness and preference on movie sales. *European Journal of Marketing*, *50* (5/6), 1024–1046.
Mostard, J., Teunter, R., & Koster, R. d. (2011). Forecasting demand for single-period products: A case study in the apparel industry. *European Journal of Operational Research*, *211* (1), 139–147.
Nguyen, H. T., & Chaudhuri, M. (2019). Making new products go viral and succeed. *International Journal of Research in Marketing*, *36* (1), 39–62.
Norton, J. A., & Bass, F. M. (1987). A diffusion theory model of adoption and substitution for successive generations of high-technology products. *Management Science*, *33* (9), 1069–1086.
Orbach, Y., & Fruchter, G. E. (2011). Forecasting sales and product evolution: The case of the hybrid/electric car. *Technological Forecasting and Social Change*, *78* (7), 1210–1226.
Ord, J. K., Fildes, R., & Kourentzes, N. (2017). Principles of business forecasting. (2nd ed.). New York: Wessex.
Pandey, H. (2024). Hathora's bare metal journey. hathora.dev, accessed on 28/11/2024 https://tinyurl.com/4237y4ds.
Perera, H. N., Hurley, J., Fahimnia, B., & Reisi, M. (2019). The human factor in supply chain forecasting: A systematic review. *European Journal of Operational Research*, *274* (2), 574–600.
Schaer, O., Kourentzes, N., & Fildes, R. (2019). Demand forecasting with user-generated online information. *International Journal of Forecasting*, *35* (1), 197–212.

Schaer, O., Kourentzes, N., & Fildes, R. (2022). Predictive competitive intelligence with prerelease online search traffic. *Production and Operations Management*, *31* (10), 3823–3839.

Schaer, O., Kourentzes, N., & Svetunkov, I. (2024). diffusion: Forecast the diffusion of new products. R package version 0.4.0. https://cran.r-project.org/package=diffusion.

Schreiner, T. F., Mandler, T., Van Heerde, H. J., & Haiduk, C. (2025). Dynamics of prerelease consumer buzz: Driving communication, search, and participation for market performance. *Journal of the Academy of Marketing Science*, *53*, 1261–1281.

Schroll, R., & Grohs, R. (2019). Uncertainty in prerelease advertising. *Journal of Advertising*, *48* (2), 167–180.

Seifert, M., Siemsen, E., Hadida, A. L., & Eisingerich, A. B. (2015). Effective judgmental forecasting in the context of fashion products. *Journal of Operations Management*, *36*, 33–45.

Situmeang, F. B. I., Leenders, M. A. A. M., & Wijnberg, N. M. (2014). The good, the bad and the variable: How evaluations of past editions influence the success of sequels. *European Journal of Marketing*, *48* (7/8), 1466–1486.

Skenderi, G., Joppi, C., Denitto, M., & Cristani, M. (2024). Well googled is half done: Multimodal forecasting of new fashion product sales with image-based google trends. *Journal of Forecasting*, *43* (6), 1982–1997.

Sood, S., & Drèze, X. (2006). Brand extensions of experiential goods: Movie sequel evaluations. *Journal of Consumer Research*, *33* (3), 352–360.

Spann, M., Fischer, M., & Tellis, G. J. (2015). Skimming or penetration? Strategic dynamic pricing for new products. *Marketing Science*, *34* (2), 235–249.

Srinivasan, S., Rutz, O. J., & Pauwels, K. (2016). Paths to and off purchase: Quantifying the impact of traditional marketing and online consumer activity. *Journal of the Academy of Marketing Science*, *44* (4), 440–453.

Stephen, A. T., & Galak, J. (2012). The effects of traditional and social earned media on sales: A study of a microlending marketplace. *Journal of Marketing Research*, *49* (5), 624–639.

Talukdar, D., Sudhir, K., & Ainslie, A. (2002). Investigating new product diffusion across products and countries. *Marketing Science*, *21* (1), 97–114.

Tassi, P. (2016). Why 'overwatch' hype is different than all other hype. Forbes.com, accessed on 24/01/2025. https://tinyurl.com/mrynzxzd.

Team, R. C. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Toubia, O., Goldenberg, J., & Garcia, R. (2014). Improving penetration forecasts using social interactions data. *Management Science*, *60* (12), 3049–3066.

Trusov, M., Rand, W., & Joshi, Y. V. (2013). Improving prelaunch diffusion forecasts: Using synthetic networks as simulated priors. *Journal of Marketing Research*, *50* (6), 675–690.

Tyebjee, T. T. (1987). Behavioral biases in new product forecasting. *International Journal of Forecasting*, *3* (3), 393–404.

van der Lans, R., van Bruggen, G., Eliashberg, J., & Wierenga, B. (2010). A viral branching model for predicting the spread of electronic word of mouth. *Marketing Science*, *29* (2), 348–365.

Van Steenbergen, R. M., & Mes, M. R. K. (2020). Forecasting demand profiles of new products. *Decision Support Systems*, *139*, 113401.

Wang, F., Zhang, Y., Li, X., & Zhu, H. (2010). Why do moviegoers go to the theater? The role of prerelease media publicity and online word of mouth in driving moviegoing behavior. *Journal of Interactive Advertising*, *11* (1), 50–62.

Welden, R., Haenlein, M., Hewett, K., Smith, K. M., & Hulland, J. (2025). Quest for insights: leveraging data from the video game ecosystem in marketing. *Journal of the Academy of Marketing Science*, *53*, 1135–1156.

Xiong, G., & Bharadwaj, S. (2014). Prerelease buzz evolution patterns and new product performance. *Marketing Science*, *33* (3), 401–421.

Yi, J., Lee, Y., & Kim, S.-H. (2019). Determinants of growth and decline in mobile game diffusion. *Journal of Business Research*, *99*, 363–372.

Yin-Poole, W. (2024). Remedy reveals how many copies control 2 and firebreak need for 100% return on investment. IGN.com, accessed on 28/11/2024. https://tinyurl.com/mwayhjud.

Zhang, M., Ahn, H.-S., & Uichanco, J. (2022). Data-driven pricing for a new product. *Operations Research*, *70* (2), 847–866.

Zhong, N., & Schweidel, D. A. (2020). Capturing changes in social media content: A multiple latent changepoint topic model. *Marketing Science*, *39* (4), 827–846.