Received: 30 August 2023 | Accepted: 12 October 2023

DOI: 10.1111/kykl.12363

ORIGINAL ARTICLE

KYKLOS WILEY

Income distribution and nudity on social media: Attention economics of Instagram stars

Sophia Gaenssle

Erasmus School of History, Culture and Communication, Erasmus University Rotterdam, Rotterdam, Netherlands

Correspondence
Sophia Gaenssle, Erasmus School of History, Culture and Communication, Erasmus University Rotterdam, Rotterdam, Netherlands.
Email: gaenssle@eshcc.eur.nl

Funding information
Ilmenau University of Technology

Abstract

Social media stars gain star-status with uploads on social media pages like YouTube, TikTok, or Instagram. One of the most popular platforms is "Instagram" owned by Meta/Facebook. The growing social, cultural, and economic power of so-called influencers raises questions about key drivers of success and, moreover, distribution of income on social media platforms. Instagram has been accused of strategically favoring images with nude content. In order to shed light on this socio-critical aspect, this paper examines the following research questions: Does body exposure drive income success on Instagram? Is there a difference between male and female content in this regard? This paper empirically analyzes 500 top Instagram stars within the categories (1) fashion and beauty, (2) fitness and sports, (3) music, (4) photo and arts, and (5) food and vegan. The data provide information on popularity, posting behavior, and price estimates per post. Using hybrid regression models, the results show indeed positive impact of body exposure on monetary success. Accounts with high level of body exposure achieve higher prices and advertising revenues than accounts with less nudity, regardless of the gender. Regarding gender differences, male content achieves on average higher advertising prices, whereas female accounts provide more branded content and eventually achieve higher advertising revenues.

1 | INTRODUCTION

Popular content providers on social media, so-called influencers, represent a novel star-type of the digital era. The fame of social media stars (SMS) is native to social media platforms (Marwick, 2016). With self-produced content, they build their audiences on pages like YouTube, TikTok, Instagram, and Twitch. In contrast to stars of traditional media, there are no gatekeepers who manage audience access (Budzinski & Gaenssle, 2020; Gräve, 2017). One of the main tasks of traditional gatekeepers is the promotion of content and generating audience attention. Content providers on Instagram & Co are able to control upload frequencies and audience access themselves, without middlemen like editors or producers. Digital developments decrease specific technological barriers to enter creative industries, such as high-quality cameras to take pictures or shoot videos and the possibility to make them publicly available on your social media page. So, from a technological point of view, it is possible to create content and put it online for the world to see, yet it is (still) very difficult to gain consumer attention. The process of audience building can be explained using attention economics, the concept of attention as a scarce resource (for an overview on attention economics, see Taylor & Greg, 2014). Because attention is scarce, it is difficult to gain even initial consumer access. To stand out and draw attention, nudity is a common way and the idea of "sex sells" has been subject to various studies (Wirtz et al., 2018). Nudity in media is not only of societal and ethical interest but also from an economic point of view. This paper empirically studies if body exposure is beneficial for Instagram success and how income is distributed among creators. Therefore, I raise the following research questions: Does body exposure drive income success on Instagram? Is there a difference between male and female content in this regard?

The service "Instagram" owned by Meta/Facebook is one of the most popular platforms and is especially designed to upload picture contents. I empirically analyze 500 top Instagram stars within the categories (1) fashion and beauty, (2) fitness and sports, (3) music, (4) photo and arts, and (5) food and vegan. The panel data set consists of 100 stars within each category over an observation period of approximately 5 months (see Section 3.1 for details on sampling). The data set contains information on popularity measurements, such as followers, likes and comments, and most importantly, price estimates per post. Because Instagram stars are not paid by the platform, but mainly by advertisers for promotion of their products (see Section 2.2 for details on social media income), the estimated price per upload is a valid proxy for economic success. The number of branded posts in combination with estimated picture prices allows a calculation of advertising revenue these stars generate. Using panel regression estimations, I statistically analyze the influence of body exposure, gender, and popularity measurements on income. The results show that body exposure positively influences advertising income, whereas impact of gender is more complex and needs careful interpretation.

The paper is structured as follows. Section 2 provides an overview on theoretical aspects such as attention economics, the role of stars, and the revenue models in social media ecosystems. Section 3 contains information on the data set, sample, and variables. In the following Section 4, the data are empirically analyzed and discussed. Section 5 gives concluding remarks, limitations, and implications.

2 | THEORETICAL BACKGROUND ON SOCIAL MEDIA STARS AND INCOME

2.1 | Stars in attention economics

The concept of limited consumer attention is not a new phenomenon of the digital age, but it excessively increased because of its developments. Simon (1971) first pointed out the problem of scarcity in an information-rich society, because information "consumes the attention of its recipients" (Simon, 1971, p. 40). With the shift to digital economy, the importance of information drastically increased, because former tangible goods could be digitized. Shapiro and Varian (1999) explain (way ahead of their time) that information goods can be all sorts of goods that can be digitized, such as films or books. The rise in available information shifted the focus; not information access, but information overload becomes prevalent in modern society (Shapiro & Varian, 1999; Taylor & Greg, 2014). Limited attention needs to be efficiently allocated dynamically over different sources and platforms (Che & Mierendorff, 2019). Therefore, limited consumer attention becomes the de facto barrier to entry in a lot of digital markets, among them social media platforms.

Budzinski and Gaenssle (2020) apply attention economics to the field of social media. This paper extends their work and theoretical idea of *audience building*. Audience building is separated into two economic effects, which take place simultaneously but can be distinguished analytically: audience *attraction* and audience *maintenance*. The former translates into first contact and creating initial attention—an eye-catcher and first connection with consumers. A common way to gain consumer attention is the paradigm of "sex sells." The exposure to sexual appeals for advertising, brand recognition/memory, and purchase intentions has been subject to many scientific studies (for detailed meta-analyses, see Wirtz et al., 2018, or Lull & Bushman, 2015). Nudity and ads with sexual appeal produce mixed results and disagreement among researchers regarding their effects on buying intentions of consumers; nudity not always increases sales (inter alia, Dudley, 1999; Mittal & Lassar, 2000; Severn et al., 1990). This study seeks to analyze the influence of *body exposure* and its influence on Instagram success, regarding the star income. Franck (2019) analyzes the "economy of attention" and refers to celebrity income as "attention income." In the following, nudity serves as variable of audience attraction and means to increase attention income. As the analysis progresses, the distribution of income is examined in more detail and differences between groups (especially with regard to gender) become clear.

The second aspect of audience building is audience maintenance. For sustainable success, it is not only necessary to generate one-time attention but to keep followers on board. The concept of consumption capital and "building of taste" can be applied here (Stigler & Becker, 1977). According to the superstar theory[1] of Adler (1985, 2006), there are three ways to acquire consumption capital:

1. Direct consumption of content: When consumers are exposed to content they build specific knowledge, for example, about the star, the type of content, and all other details and knowledge that is transmitted. A food account, for instance, can serve as information on health and cooking advice, and additionally, the fans get used to the content provider and her type of presentation and communication. As Adler puts it in his seminal paper on superstars, "The more you know, the more you enjoy" (Adler, 1985, pp. 208–209). The more videos are available on an Instagram account, the better the possibility to acquire consumption capital. Fans can get to know the content provider, scroll through her page, and build specific knowledge. Therefore, in this study, the *number of total posts* on the account serves as proxy for the possibility to acquire consumption capital.
2. Communication about content (communality effects): Consumers not only build specific knowledge and derive utility from consuming the content itself, but also through communications with others (commonality effect), that is, exchanging information with others and learning even more. Consumption capital, thus, drives bandwagon effects (Leibenstein, 1950). The communicational aspect gained importance in the digital environment. Cost of communication drastically lowered on the demand and the supply side (Gaenssle & Budzinski, 2021). Fans can easily comment on posts of their favorite stars and connect among each other; emojis and likes even simplify this process and further reduce costs of communication. It is very easy to find like-minded people and forward content to friends and family. This leads to considerable network effects (Katz & Shapiro, 1985, 1994). Jung and Nüesch (2019) find that popularity serves not only as a quality signal, but the mere number of historical views on a YouTube video generates utility itself. Consumers perceive these videos as popular; they can share and talk about it. There are different popularity indicators; the most common ones are views, followers, likes, and

[1]Within the discussion of economics of superstars, first introduced by Rosen (1981), the aspect of talent is of concern. However, objective measures of talent have proven to be a considerable challenge in empirics (Franck & Nüesch, 2012; Krueger, 2005). This article therefore focuses network effects and aspects later introduced to the economic theory on superstars by (Adler, 1985).

comments (Burgess & Green, 2018). To operationalize the popularity concept in this paper, I use the total number of *followers* and the average of recent *likes* and *post engagement*.

3. Media coverage (availability): According to Adler (1985, 2006), the cost of consumption arises from (1) actual cost of consumption, that is, watching videos, looking at pictures, and reading captions, and (2) search cost of finding suitable content and conversation partners. If content availability is necessary to lower cost of consumption and drive social media success, international strategies have to be taken into account. Therefore, I calculate the concentration of the audience regarding their country of origin, spoken language, and interests using Herfindahl–Hirschman Indices (*hhilocation*, *hhilanguage*, *hhiinterest*). The concentration measures serve as a proxy for availability and spread of the content.

Both theoretical parts of audience building deal with scarce attention; attraction and maintenance need the consumers' time and focus. The second part, however, pertains to superstar theory, which is very well researched. A lot of empirical evidence can be found focusing on ideas of Adler and his concept of snowballing into superstardom.[2] This paper aims to contribute to filling the research gap and analyze the first part, the initial attraction of audience attention, using body exposure as a proxy. Furthermore, it shows the connection between attention and star income as well as income distortions (despite high attention and popularity).

2.2 | Sources of income for social media stars

There are different ways to generate revenue as a content provider in the social media system. In the following, a general overview shows the different possibilities to earn money on different platforms. Not all options are available on every platform. The list gives a systematic overview and some examples (in *italic*).

Types of platform immanent payments:

i. Share of advertising revenue: This is the case, inter alia, on the video platform *YouTube*. Content providers agree to embed advertising before, during, or/and after their videos and get paid per view—the more ads are consumed, the better the payment.
ii. Direct donations (tips): It is possible to directly send money to the content providers, that is, tip them. This is (among other payment methods) possible on video and gaming platform *Twitch*.
iii. A payment barrier to access content: The content provider can establish a paywall, for instance, a monthly flat rate (all you can eat) system or pay-per-view (à la carte) system. This is, for instance, possible on the platform *Patreon*, where you can support your favorite artist and access the content on a subscription base.

Types of external payments:

iv. Own sponsored content or own brands: With growing reach content providers often start their own business or sell and promote their own products. A baker for example sells baking equipment, and a makeup artist owns makeup products. This is broadly used on *WeChat* in connection with *Taobao* shops in China; content providers have their own shop and products connected to the social media page. New options of *Instagram* enable similar functions.

[2] A literature overview on economics of superstars can be found in Budzinski and Gaenssle (2020). Empirical papers testing superstar concepts: inter alia, Budzinski and Pannicke (2017); Bryson et al. (2014); Candela et al. (2016); Crain and Tollison (2002); Ehrmann et al. (2009); Filimon et al. (2011); E. Franck and Nüesch (2007, 2008, 2012); Giles (2006); Hofmann and Opitz (2019); Jung and Nüesch (2019); Lehmann and Schulze (2008); Lucifora and Simmons (2003); Meiseberg (2014); Salganik et al. (2006), Torgler et al. (2008).

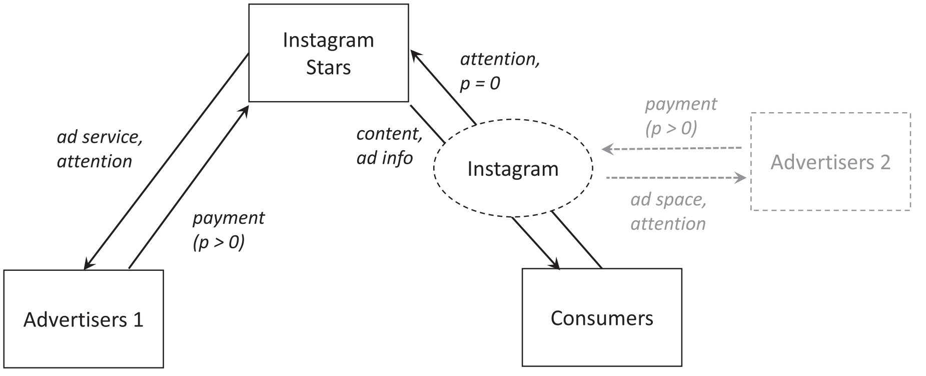

FIGURE 1 Ecosystem of platforms: Instagram and Instagram stars.

v. Commission and affiliated links: It is possible to embed affiliate links (e.g., *Amazon* cooperates with many retailers) or use discount codes to get a commission for sales. The latter is also popular on the social media platform *Instagram*.

vi. Ad financed (or sponsored) content: This can start very small with product placements and free samples. Content providers include products in their content and promote them. Very successful social media stars can earn vast amounts of money to include ads in their posts. This payment strategy is predominant on *Instagram*.

Version (vi) is used within this study. It is a valid proxy for income on Instagram because most of the top stars use ad financing. Figure 1 shows the Instagram ecosystem for deeper understanding. There are two advertising parties, those directly paying Instagram for placing ads in user chronicles (Advertisers 2) and those cooperating with Instagram stars (Advertisers 1). Even though Instagram itself cooperates with Advertisers 2, they do not act as intermediaries and give Instagram stars the option to monetize their content though the platform. In comparison with YouTube, Instagram stars cannot embed ads in their content through the platform and get "paid per view." Instagram does not offer platform immanent payments. Therefore, Instagram stars cooperate directly with advertisers and not though the platform. This is why option (vi) is prevalent and not (i).

The Instagram stars face two demand groups: (i) Consumers (though Instagram) and (ii) Advertisers 1. The more consumers a star can reach, the stronger the demand effect on advertisers (positive indirect network effect). The stars serve two distinct demand groups; there are indirect effects between those groups that they internalize through content provision. Hence, from an economic point of view, the stars function as intermediaries in a two-sided market and platforms themselves (inter alia on platform economics: Armstrong, 2006; Evans & Schmalensee, 2007; Rochet & Tirole, 2003, 2006). The figure illustrates the platform ecosystem. The stars serve more than one demand group and Instagram itself too.[3] Essential for this paper are the streams of revenue from Advertisers 1 to Instagram stars. These revenues serve as income proxy (more details on the variables follow within the empirical analysis).

[3]Within this consideration, Instagram itself takes the position between the stars (or other content providers respectively) and consumers. Instagram functions as a gatekeeper and filter. Not all posts reach the entire audience. Instagram can pre-select according to preferences (and other settings) and has algorithmic recommendations for consumers based on personalized data (Budzinski et al., 2021). As such, and from the star's point of view, it acts as a filter, managing which types of content reach the target group. Gaenssle and Budzinski (2021) point out the importance of algorithm management skills for content providers on social media and provide further insights regarding this topic.

3 | DATA AND VARIABLES

3.1 | Sample

The data are retrieved from Heepsy.com and Instagram. Heepsy primarily provides data for companies, who want to find influencers for ads on Instagram. Therefore, the service includes prices per post. To retrieve data from Heepsy, a professional account with payment is needed. The funding of the paper allowed a time series of 6 months and 500 influencers, with observations approximately every 10 days (translating into 15 observations per account on average). Consequently, I have an unbalanced data set (unbalanced due to availability issues and cleaning of duplicates), with 7362 observations in total.

It is possible to select the Instagram star according to specific settings on Heepsy. Only accounts that stated their mail address on Instagram are included in the sample. Adding a mail address to the profile can be interpreted as interest in further contact, apart from direct messages within Instagram. It can be expected that stars with email address want to be contacted and show business interest. Five popular but distinct categories are selected: (1) beauty & fashion (Fashion); (2) fitness and yoga (Fitness); (3) food and health (Food); (4) music (Music); and (5) photo and arts (Photo). The choice was influenced by the popularity of the categories and their distinctiveness (Hypeauditor, 2020). Some categories are broadly overlapping (e.g., sports and fitness). In case of doubt and considering double sampling, only the more popular one was chosen for the final sample. Within each category, the top 100 accounts are sampled. Table 1 shows the final sample and observations per category. The data varied depending on the updates of the page Heepsy.com.

3.2 | Dependent variable

The following section explains the nature of the dependent variable. It does not serve as a theoretical model for the study but aims to explain the economic nature of the variable and its characteristics. This is done in a formal way to clarify some specific features of the industry and revenue estimations. The dependent variable for this study is the star income of advertisement, that is, the advertising revenue. This serves as a proxy for their income through social media activity. Conventionally, the revenue is calculated by price times quantity. Within this platform ecosystem (explained in Section 2.2), the Instagram star faces two demand groups: (i) content consumers and (ii) advertisers. The star produces content to create a fan base and needs to upload on a regular basis. Empirical evidence shows a U-shaped relation between popularity of content and uploading behavior. Content providers need to upload regularly and show activity, but they should not overstrain their audience (i.e., spamming) and provide too much content, creating information overload (Budzinski & Gaenssle, 2020). Eventually, most Instagram stars provide not branded

TABLE 1 Overview observations per category.

IG category	Freq.	Percent	Cum.
Fashion	1501	20.39	20.39
Fitness	1501	20.39	40.78
Food	1510	20.51	61.29
Music	1504	20.43	81.72
Photo	1346	18.28	100
Total	7362	100	

content q^c to keep their audience entertained and branded[4] content, which is paid for by advertisers q^a. This translates into the following equation for the revenue R.

$$R = pq = p^a q^a + p^c q^c \tag{1}$$

Content consumers pay attention and time, but no monetary price within the Instagram system. Therefore, $p^c = 0$. However, advertisers pay a monetary price $p^a > 0$. The revenue generated on the side of advertisers also depends on the number of consumers on the other side of the market, that is, advertisers are interested in reaching as many consumers as possible via the intermediary Instagram star. Therefore, p^a depends negatively not only on the quantity of advertised pictures q^a but also on the quantity of non-branded pictures q^c, weighted by the indirect network effect g between the market sides (see Equation 2). Therefore, the revenue R generated by Instagram stars is generated on the advertising market with an indirect network effect from consumers (Equation 3)

$$p^a = 1 - q^a + gq^c \tag{2}$$

$$R = p^a q^a = (1 - q^a + gq^c) q^a \tag{3}$$

Advertisers can place ads with influencers and pay per picture. In economic terms, the pictures are *heterogenous goods* and advertisers pay heterogeneous prices for *each singular* picture. Thus, the influencer promotes their product in one paid picture. The prices are individual and distinct for every upload. Therefore, the income for Instagram stars consists of the sum of advertised images.

$$R_n = (1 - q^a + gq^c) q_1^a + (1 - q^a + gq^c) q_2^a + \ldots + (1 - q^a + gq^c) q_n^a \tag{4}$$

$$R_n = \sum_{k=1}^{n} (1 - q^a + gq^c) q_k^a \tag{5}$$

Heepsy.com offers market-based information[5] and calculates the average price that an advertiser would have to pay to be advertised in *one* image, that is, the result of p^a including the indirect network effects. Moreover, the posting frequency and the percentage of branded posts allow a calculation of q^a (the quantity of advertised posts in that period). Within the data set, I can calculate the average ad revenue $\bar{R}$ based on the price p^a and the posting frequency q^a (theoretically represented in Equation 6).

$$\overline{R_n} = \frac{1}{n} \sum_{k=1}^{n} (1 - q^a + gq^c) q_k^a \tag{6}$$

For instance, the price for makeup artist Patrickstarrr end of January 2020 was USD 6.100 (p^a) per picture. Hence, an advertiser who wanted to promote a product with Patrickstarrr was expected to pay this amount for one single image. He had 4,782,780 followers, and his pictures reached on average 55.246 likes and 221 comments. The star posted on average 3.3 pictures per week, and 42% of the content was branded, that is, ~1.4 pictures. Therefore, his average ad revenue is *price per picture* times *percentage of branded posts* times *posting frequency* equals to 8454.60 USD per week.

[4]Budzinski and Gaenssle (2020) do not find empirical evidence that branded posts are overall less popular than unbranded posts.

[5]Further details of data collection etc. are business secrets and not publicly available. This information can be obtained only through personal contact. As correctness and reliability is essential for Heepsy to calculate accurate costs for their customers, it can be expected that the proxy is valid.

Table 2 gives an overview of p^a; the superstar prices have a mean of 16,467.84 USD per picture, but there are also big differences between accounts within the sample with a standard deviation of 28,399.38. The weekly posting frequency q^a of branded pictures is displayed in Table 3. On average, the content creators post 2.5 branded pictures per week.

Table 4 shows the average ad revenues calculated as described above. The total mean over the whole sample is 112,943 USD. There are also periods where stars did not post branded pictures and the revenue is 0. Stars in the category Music generate the highest revenues, followed by Fashion and Fitness. Actors within Food and Photo cannot achieve the high levels of the first three categories.

Because the data are right skewed and, thus, residuals have a skewed distribution, the data are transformed to obtain symmetrically distributed values and operate linear equation estimations. The logarithm of the income proxy provides a good fit, compared with the log-normal distribution. Figure 2 illustrates the distribution against a normal distribution. Eventually, the logged values $ln\bar{R}$ are the dependent variable and the income proxy for the following analysis.

TABLE 2 Price per Pic p^a per category (USD) (January–June 2020).

IG category	Freq.	Mean	Std. dev.	Min	Max
Fashion	1501	21,052.21	32,733.77	145	345,000
Fitness	1501	15,513.82	28,068.31	180	300,000
Food	1510	6618.50	8640.11	18.5	138,500
Music	1504	30,170.94	38,979.36	700	345,000
Photo	1346	8157.23	12,095.32	26	107,000
Total	7362	16,467.84	28,399.38	18.5	345,000

TABLE 3 Posting frequency branded picture per week q^a per category.

IG category	Freq.	Mean	Std. dev.	Min	Max
Fashion	1218	2.40	4.00	0	37
Fitness	1410	2.35	3.82	0	46
Food	1224	2.51	12.24	0	409
Music	1411	1.94	2.87	0	29
Photo	1261	3.37	14.77	0	210
Total	6524	2.50	8.85	0	408.70

Note: The total number of observations is reduced because of missing information on *percentage of branded posts*.

TABLE 4 Advertising revenue $\bar{R}$ per week per category (USD).

IG category	Freq.	Mean	Std. dev.	Min	Max
Fashion	1218	30,628.84	65,435.32	0	599,841
Fitness	1410	21,069.90	52,958.52	0	892,500
Food	1224	10,445.97	20,156.91	0	282,750
Music	1411	41,304.79	74,603.25	0	982,080
Photo	1261	9493.87	20,262.13	0	276,375
Total	6524	23,000.18	53,964.73	0	982,080

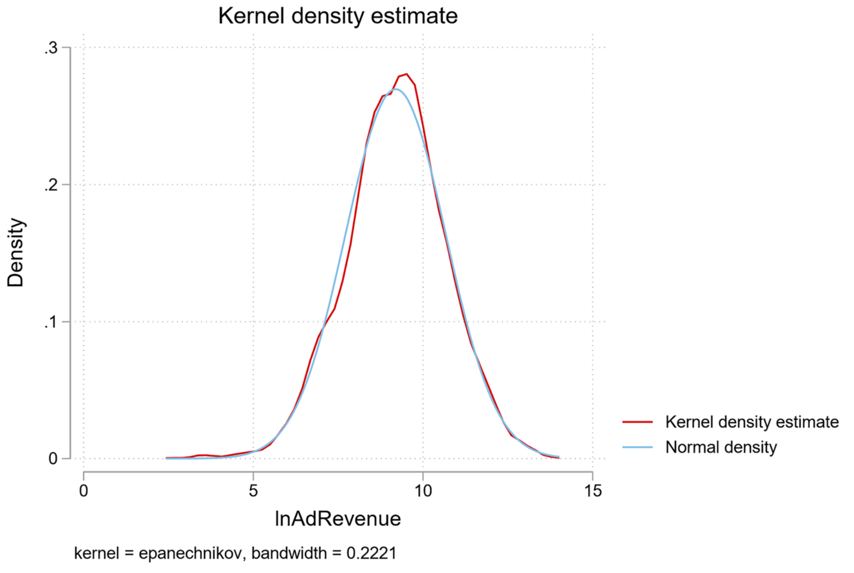

FIGURE 2 Overview distribution: Logged values advertising revenue $\bar{R}$. [Colour figure can be viewed at wileyonlinelibrary.com]

3.3 | Independent variables and controls

I use two sets of independent variables: the ones for audience attraction and those for audience maintenance. Body exposure is used as a proxy for audience attraction. It is a time-invariant variable, because the coding is very extensive, that is, the process was only performed once. The last 12 pictures of the account and their respective degree of nudity are analyzed.[6] Different types of nudity can be defined theoretically, as in everyday dresses ("demure"), mini-skirts or shorts ("suggestive"), bathing suits/shorts ("partially clad"), and without clothes ("nude") (Reichert, 2002; Wirtz et al., 2018). Instagram's guidelines restrict nudity to a certain degree.[7] For this study, the category "partially clad" is defined more closely into the degree of nudity. Therefore, the degree of nudity within the last active pictures posted (12 pics per account) is analyzed. If a picture shows 50% or more naked skin (excluding portrait pictures), the body exposure is coded as 1. If there is less nudity in the picture, it is coded as 0. To account for the degree of suggestiveness, if 50% or more focus lie on "dressed" primary sexual characteristics (breasts, bottom, genitals), this is also coded as 1 (see examples in appendix Figure A1). I operationalize the degree of nudity on the account by summing up the "nude pictures," that is, accounts with body exposure = 12 have a very high degree of nudity, accounts with body exposure = 0 have no nude pictures. The percentage of body exposure pictures is calculated for every account, to improve transparency and understanding in the descriptive analysis. Because the coding process is sensitive to subjective perception, a second person evaluated the pictures for inter-coder reliability. Moreover, to ensure intra-coder reliability, a second test was performed, that is, the same coder rechecked the pictures a few months later (Figure A1).

[6]Heepsy uses 12 pictures for their analyses and displays them for customers. Those 12 pictures are also used for further manual coding in this paper. Instagram Stories are not part of the analysis.

[7]Community guidelines by Instagram (as in spring 2021): "Post photos and videos that are appropriate for a diverse audience: We know that there are times when people might want to share nude images that are artistic or creative in nature, but for a variety of reasons, we don't allow nudity on Instagram. This includes photos, videos, and some digitally-created content that show sexual intercourse, genitals, and close-ups of fully-nude buttocks. It also includes some photos of female nipples, but photos in the context of breastfeeding, birth giving and after-birth moments, health-related situations (for example, post-mastectomy, breast cancer awareness or gender confirmation surgery) or an act of protest are allowed. Nudity in photos of paintings and sculptures is OK, too." Facebook (2021)

The second set of independent variables are the ones for *audience maintenance* focusing on (i) posting behavior, (ii) popularity, and (iii) availability/audience measurements. They are summarized in Table 5 and vary over time, that is, between observations.

i. Posting behavior: The *Numberofposts* is the number of total uploads on the account, that is, total content available. The mean number of posts is 2968.96 pictures with a standard deviation of 4278.14; there are big differences between accounts. Some post a lot more (higher frequency) than others. The age of the account is unknown; hence, it is not possible to control how much time the Instagram star spends to upload content.
ii. Popularity indicator: The number of total *Followers* shows the fan base of the star and the regular audience. Because popular accounts were chosen, the overall mean is 6.5 million followers. The most followed has 65.8 million and the least followed 1.4 million. The variable *Postengagement* is the sum of all follower engagement with the star (likes, comments, shared, etc.) relative to the number of followers. It turned out to be superior to "comments" in the analysis with higher explanatory power. The average of the sample is 2.58, but the maximum is 23%. The *likes* an account receives is the best proxy for popularity within the analysis, with a mean of 162.4 thousand and a standard deviation of 258.0 thousand.
iii. Audience measurement and availability: I calculated the concentration measurement HHI (Herfindahl–Hirschman Index) of audience location, language, and interests. This can be used to control how concentrated the audience is, for example, in how many countries the star is represented, how many languages the audience

TABLE 5 Overview of independent variables.

Variable		Mean	Std. dev.	Min	Max	Obs.
Numberofposts	Overall	2968.96	4278.14	12.00	57,048.00	$N = 7362$
	Between		4174.57	17.00	56,149.19	$n = 500$
	Within		141.28	−79.66	7117.15	
Followers	Overall	6,474,005	6,587,006	1,436,180	65,798,946	$N = 7362$
	Between		6,410,906	1,455,425	61,600,000	$n = 500$
	Within		406,500	1,446,296	10,700,000	
Postengagement	Overall	2.58	2.76	0.01	23.00	$N = 7362$
	Between		2.70	0.03	17.36	$n = 500$
	Within		0.86	−5.89	13.95	
Likes	Overall	162,385.10	258,043.00	91.00	3,234,545.00	$N = 7362$
	Between		241,077.60	455.80	2,108,876.00	$n = 500$
	Within		92,460.28	−1,121,365.00	1,475,279.00	
HHI audience location	Overall	2903.52	2468.20	319.00	10,000.00	$N = 7158$
	Between		2477.17	319.00	10,000.00	$n = 466$
	Within		0.00	2903.52	2903.52	
HHI audience language	Overall	5437.85	2022.50	1339.00	9605.00	$N = 7158$
	Between		2037.15	1339.00	9605.00	$n = 466$
	Within		0.00	5437.85	5437.85	
HHI audience interests	Overall	1262.93	527.95	538.00	4407.00	$N = 7158$
	Between		529.90	538.00	4407.00	$n = 466$
	Within		0.00	1262.93	1262.93	

Note: Audience information was not available for all accounts. Therefore, the *n* is reduced to 466 accounts regarding these variables (*HHI Audience Location*, *Language*, and *Interests*).

speaks, and if their interests are diverse or homogenous. Because of the formula of the HHI, the maximum can be 10,000 (100% concentrated), which is the maximum for *HHI Audience Location* (all audience in one country), and the mean is 2903.52. *HHI Audience Language* is on average higher concentrated with a mean of 5437.85 and *HHI Audience Interests* is least concentrated with 1262.93.

More complex is the information on the type of content on the account. The measurement of gender refers to the type of content and what the focus of the account is, that is, is there a focus on male, female, or no specific focus on the account? Therefore, all accounts are analyzed individually. For decades, researchers have carefully studied the terms sex (biology/physiology) and gender (social/sexual identity, created impression) (Lorber, 1996, 2005). A researcher can only observe the created impression or social perception. An external identification of the gender is only possible as "read as" female or male, according to visible features. I coded five different variables: (1) female (clearly female features, one or more women); (2) male (clearly male features, one or more men); (3) mixed (clearly male and female, if more than one person in the picture); (4) ambivalent (sexual characteristics recognizable, but not clearly attributable to a man or a woman, e.g., transvestite or similar); and (5) no identification (no characteristics visible). This goes hand in hand with the body exposure analysis; 12 pictures per account are analyzed and coded. If all 12 pictures show a woman, this is coded as 12-0-0-0-0, and if there are 6 pictures of women and 6 pictures of men, this is coded as 6-6-0-0-0 (see example in Appendix A). Thus, it is possible to implement a proxy—the degree of female or male pictures of every account. On average, there is more female content in the sample with 51.88% versus 21.6% male content. The rest is even lower with 13.05% mixed, 12.8% no identification, and 0.27% ambivalent.

4 | EMPIRICAL ANALYSIS OF INSTAGRAM STARS

4.1 | Results descriptive analysis and mean comparison

Within the sample, it is interesting to study the top positions and means to gain first insights into the data and success factors of the "top crowd." The analysis of the top 20 over the whole sample (Table 6) and the top 10 (Table 7) within each category already shows some first connections and impressions. They are ordered by the dependent variable price per picture—the top accounts according to mean price per picture over all observation periods. Seven of the overall top 20 are Music accounts, followed by Fashion and Fitness. There is one Photo account in the top 20, but no Food account. Most of the content shows women with 54.58%. Given that the sample was created by collecting top accounts within the categories, this shows that the majority of the top accounts (among the already most successful ones) show female content. It implies that female content is more popular, or at least most frequently represented. Within the top 20 accounts with mixed gender make up 24% and with clear focus on male content 19%. The average body exposure is 37%—so less than half of the pictures on the accounts are on average nude pictures. But there are accounts with 0 and accounts with 100% body exposure. Therefore, nude photos are not essential for success within the sample, but they are nevertheless common.

TABLE 6 Gender and body exposure for the top 20 overall (mean picprice).

Categories		Gender (%)		Body exposure (%)
Fashion	7	Female	53	37 (min 0, max 100)
Fitness	5	Male	19	
Food	0	Mixed	24	
Music	7	Ambivalent	0	
Photo	1	No identification	4	

TABLE 7 Gender and body exposure for the top 10 per category (mean price per picture).

Category	Gender (%)					BE (%)
	Female	Male	Mixed	Ambiv.	No ident.	
Fashion	64	16	20	0	0	37
Fitness	38	34	24	0	4	40
Food	36	42	13	0	10	5
Music	44	29	19	0	8	24
Photo	55	26	15	0	4	16

The next step is to look deeper into the respective categories. Within the top 10 per category, the majority of contents display women, followed by men. Female photos are most common in Fashion, followed by Photo, Music, Fitness, and Food. Male contents are represented most among food accounts, followed by Fitness, Music, Photo, and Fashion. Mixed contents vary between 24% and 13% over all categories, whereas there is no ambivalent account within the top 10 per category. The category Food shows the highest proportion of "No Identification," evidently with food pictures. Expectantly, the body exposure is also lowest in Food. The highest body exposure is in Fitness with 40%, followed by Fashion with 37%. Overall, the top 10 per category show that female and male content is most frequent and body exposure is most prevalent in Fashion and Fitness.

When analyzing body exposure, the question of gender is ubiquitous. The differences between male and female accounts are of interest, because they make up a majority of the overall sample. There are some accounts in the total sample, which do not have a clear focus on either male or female content. I exclude those accounts to study whether there is a difference between accounts with clear focus on either male content or female content. I restrict the observations to those accounts with "clear focus" (more than 50% male/female content). This excludes accounts with diverse uploads, that is, mixed accounts with females and males, non-human accounts (like memes, landscapes, or pets), and non-focus accounts (a little bit of everything). This leaves 382 accounts. With this sample, I perform mean comparison tests (t-tests) to see if there are differences between women and men.

i. *Body Exposure*: The results show that accounts with focus on female contents have significantly higher body exposure than accounts with focus on male contents (p-value = 0.0000). As such, women appear to show more nudity (mean men = 1.25 pics out of 12, mean women 4.3 pics out of 12).

$$\text{body exposure}_{\text{male}} \ll \text{body exposure}_{\text{female}}$$

ii. *Posting frequency*: There is a statistically significant dissimilarity in posting frequency (p-value = 0.0003). Women upload significantly more advertised content (mean men = 1.72 pics per week, mean women 2.17 pics per week).

$$q^a_{\text{male}} < q^q_{\text{female}}$$

iii. *Price per picture*: There is a difference in price per picture (p-value = 0.0017). Male contents achieve significantly higher price per picture. This is surprising, because female content makes up a large proportion of the top content. Yet, within the sample, male content is more successful in price per picture (mean men = 19,521.41 USD per pic, mean women 16,998.03 USD per pic).

$$p^a_{\text{male}} > p^a_{\text{female}}$$

iv. *Average advertising revenue*: Although female accounts achieve lower prices per picture, their revenue is significantly higher. The difference in posting frequency compensates for the price difference, so that women ultimately achieve higher ad revenues (p-value $= 0.0238$), (mean men $=$ 22,654.02 USD per week, mean women 26,209.39 USD per week).

$$\overline{R}_{\text{male}} < \overline{R}_{\text{female}}$$

To sum up, accounts with female contents show a higher degree in nudity, but they earn less per picture. They post significantly more advertised content and therefore achieve higher revenues eventually.

4.2 | Results analytical analysis and regression estimations

4.2.1 | Research question 1

The first results have already provided interesting insights. However, different sets of covariates and time effects are not considered. That is why more sophisticated methods are used for the analytical analysis of the data set. To answer the research question *Does body exposure drive income success on Instagram?*, more complex models are necessary. The influence of body exposure and gender stands in the focus of this section. Therefore, I use panel regression estimations.

The data set contains time-variant variables (followers, likes, etc.) and time-invariant variables (category, body exposure, and gender). Therefore, a standard fixed effects model does not fit the theoretical framework and research questions. Key independent variables would not be identified by a fixed effects specification. Because the influence of body exposure and gender is central, the estimation results for those variables are vital. This can only be achieved by including them using random effects. However, in random effects models, the assumption is that there is no correlation between the unobserved heterogeneity. That is why I use a hybrid model (Allison, 2009) where within effects are possible in random effects model by decomposing x_{it} in a between $\overline{x}_i$ and a cluster $x_{it} - \overline{x}_i$ component. One advantage of this approach is that it allows to test for within and between estimates equivalently (Schunck, 2013). Equation (7) formally expresses the hybrid model:

$$y_{\text{it}} = \beta_0 + \beta_1 \left(x_{\text{it}} - \overline{x}_i\right) + \beta_2 c_i + \beta_3 \overline{x}_i + \mu_i + \varepsilon_{\text{it}} \tag{7}$$

where *i* is the account and *t* is the time of observation (occasions). x_{it} varies between occasions and within clusters (*followers*, *posts*, etc.), and c_i varies only between clusters (*gender*, *category*, *be*, *HHIs*) (time invariant). μ_i is the random intercept and between error (only varies between clusters), and ε_{it} is the idiosyncratic error term (varies within clusters). β_1 gives within-effect estimate (fixed effects estimate), and β_3 estimates the between effect (Mundlak, 1978; Neuhaus & Kalbfleisch, 1998). This way, I can observe if differences across entities (body exposure) have influence on the dependent variable.

The estimation results are shown in Table 8. The dependent variable for all estimations in this table is *lnAdRevenue*, the logged values for $\overline{R}$ (as explained in Section 3.2). I estimate four models with different sets of independent variables. Because popularity factors are highly correlated (see Table A1), I integrate them individually. The first model shows results for likes; the second includes followers and post engagement. Because followers and post engagement increase the explanatory power of the model (overall R-squared: 0.3408 vs. 0.4250), I further enhance Model 2. Model 3 includes the audience information (concentration indices HHI on location, language, and interests). Because there are some missing observations for the HHIs, the n is lower in the last two models. I use generalized estimating equations, which is possible when decomposing into between and within effects and enables a less restrictive within-cluster error specification (Schunck, 2013).

TABLE 8 Hybrid models: Research question 1.

Variables	(1) lnAdrevenue (likes)	(2) lnAdrevenue (followers)	(3) lnAdrevenue (demand)	(4) lnAdrevenue (interaction)
Mean(s) of kNumberofposts	0.046***	0.053***	0.052***	0.053***
	(0.012)	(0.012)	(0.012)	(0.012)
kNumberofposts (centered)	0.109*	0.148***	0.148***	0.148***
	(0.059)	(0.051)	(0.051)	(0.051)
Mean(s) of kLikes	0.003***			
	(0.000)			
kLikes (centered)	0.002***			
	(0.000)			
Mean(s) of mFollowers		0.089***	0.088***	0.090***
		(0.011)	(0.011)	(0.010)
mFollowers (centered)		0.015	0.015	0.015
		(0.031)	(0.031)	(0.031)
Mean(s) of Postengagement		0.233***	0.232***	0.235***
		(0.020)	(0.020)	(0.020)
Post engagement % (centered)		0.234***	0.231***	0.231***
		(0.021)	(0.021)	(0.021)
Female pics	0.044**	0.023	0.016	
	(0.019)	(0.018)	(0.018)	
Male pics	0.045**	0.019	0.012	
	(0.020)	(0.019)	(0.019)	
Body exposure	0.031**	0.039***	0.039***	0.052**
	(0.015)	(0.014)	(0.014)	(0.025)
Category # body exposure				Figure 3
Category = 2, Fitness	−0.066	−0.030	−0.022	−0.128
	(0.149)	(0.145)	(0.145)	(0.192)
Category = 3, Food	−0.373**	−0.234	−0.246*	−0.231
	(0.160)	(0.148)	(0.147)	(0.178)
Category = 4, Music	0.317**	0.251**	0.232*	0.385**
	(0.138)	(0.127)	(0.130)	(0.160)
Category = 5, Photo	−0.648***	−0.549***	−0.568***	−0.527***
	(0.177)	(0.163)	(0.166)	(0.192)
HHI audience location			Incl.	Incl.
HHI audience language			Incl.	Incl.
HHI audience interests			Incl.	Incl.
Constant	8.137***	7.538***	7.566***	7.645***
	(0.215)	(0.212)	(0.311)	(0.257)
R-squared within	0.0500	0.0922	0.0915	0.0915
R-squared between	0.4143	0.5069	0.5057	0.5099

TABLE 8 (Continued)

Variables	(1) lnAdrevenue (likes)	(2) lnAdrevenue (followers)	(3) lnAdrevenue (demand)	(4) lnAdrevenue (interaction)
R-squared overall	0.3408	0.4250	0.4265	0.4321
Observations	5458	5458	5320	5320
Number of id	485	485	454	454

Note: The dependent variable for all estimations in this table is lnAdRevenue, the logged values for advertising revenues $\bar{R}$ (as explained in Section 3.2).
In Models 1 and 2: $N = 5458$ and $n = 485$; only accounts with positive revenues are included within the estimation. In Models 3 and 4: $N = 5320$ and $n = 454$; not all accounts include audience information.
Models 1 and 2 include different independent variables for popularity: Model 1 the number *Likes* and Model 2 the number of *Followers* and *Post Engagement*. Because the explanatory power of Model 3 is higher, it serves as base for the following regressions. Model 3 extends Model 2 regarding the concentration of audience. Model 4 includes an interaction term between *Category* and *Body Exposure*.
All estimation are hybrid models with a between (between different accounts) and a cluster component (within the same account). *Mean(s)* are cluster-specific means and *(centered)* are deviation scores between accounts to implement the hybrid model. Some variables are scaled in size (k in thousand, m in million). The base category is 1. Fashion (with 2. Fitness, 3. Food, 4. Music, 5. Photo). Robust standard errors are in parentheses.
***p < 0.01, **p < 0.05, and *p < 0.1.

The first main result is that body exposure has a significant positive effect on income in all four models (and within all further models on robustness, see Section 4.3). The impact of the estimated coefficient $\hat{\beta} = 0.039$ of Models 2 and 3 can be calculated by a quick approximation $e^{0.039} \approx 1.0398$ (Benoit, 2011), which means that within the data set, one increase in body exposure (one more picture coded as nude) increases the advertising revenue by 3.9%. By comparing different portfolios of accounts (some having a higher degree in body exposure than others), the estimations show that, on average, one more nude picture increased the revenue by almost 4%. Because of the temporal invariance of the variable body exposure, only inter-account estimation is feasible, precluding the determination of the marginal effect of a single additional nude image within the same account. However, the estimation illustrates that accounts with higher degree in nudity perform comparably better, therefore demonstrating that accounts exhibiting a greater degree of nudity exhibit relatively superior performance. Over the whole data set, one more picture showing naked skin increases revenues by almost 4%. Given our focus on income of superstars, the percentage change in revenue is considerably high.

Upon examination of the remaining independent variables, the following observations can be made. Expectantly, the popularity indicators are significantly positive, so is the post frequency. These results are in line with past studies (Budzinski & Gaenssle, 2020; Jung & Nüesch, 2019), which can be seen as a quality indicator and shows reliability of the data set. Regarding gender, only Model 1 holds significant results; robustness checks also show instable results. This is why these coefficients are not interpreted here but within the subsequent section.

Figure 3 illustrates the margins plot for Model 4. The vertical axis shows the level of lnAdrevenue, whereas the horizontal axis displays levels of nudity. The results do not show the marginal effect of another nude picture *within* an account, but the comparison *between* accounts. All margins are above zero and therefore significantly positive. Thus, the mean marginal impact of a nude image on the overall sample is positive, with increasingly wider confidence intervals as the level of nudity rises. First, to compare the effects within a category, the confidence intervals must be considered. If these intervals do not intersect, it indicates that there are significant differences between the units of body exposure. Within the Fitness category, there is no overlap between the confidence intervals of the first two and last two units. Therefore, it can be said that accounts with higher degree in nudity perform comparably better than accounts at a low level. Within the other categories, there seems to be no such effect and the confidence intervals intersect. Second, confidence intervals between categories can be compared. The estimates vary around the

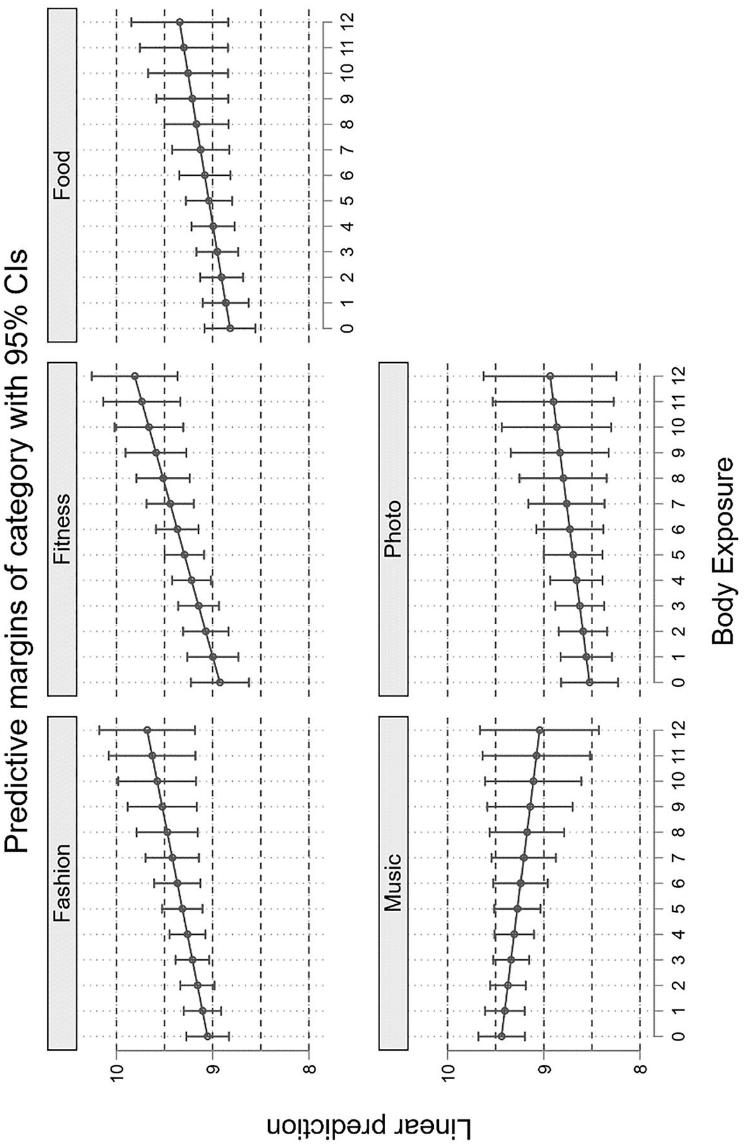

FIGURE 3 Marginal effects of Model 3: Body exposure per category.

value of 9, mostly overlapping. Considerable differences between the categories can only be found for low levels of nudity between Music, Photo, and Food. Although in the latter, two higher levels of exposure tend to generate greater income, in the category Music, higher degrees of body exposure are inversely associated with revenue. Accordingly, Music accounts with a low proportion of nude images are significantly different from those of the other two categories and perform comparably better.

Overall, there is a positive slope observable, underlining the findings that there is a positive relationship between body exposure and advertising revenues. Except for Music, the trend is "accounts with higher degree of nudity can achieve higher levels of income." However, it must be said that the analysis reveals only a trend, but it does not show systematic deviations within and between all categories. Only within the category Fitness, the effects are so strong that significant distinctions can be observed. Here, nudity is specifically beneficial. In a category that focuses on physical activity, this seems plausible. Followers can feel inspired by the activity and shape of their idols and influencers (Gaenssle & Budzinski, 2021).

Figure 3 shows the marginal effects of *Body Exposure* (x-axis) on the *lnAdRevenue* (y-axis) per *Category*. If confidence intervals overlap, there are not significant differences. Only a few observations show significant differences. Within the category Fitness accounts with high level of body exposure perform significantly better than accounts with low levels. In the category Music, accounts with low level of body exposure perform significantly better than accounts in Photo or Food.

4.2.2 | Research question 2

The analysis so far shows that body exposure has a positive impact on income from advertising activities within this sample. However, no reliable gender effects were found to answer research question 2: *Is there a difference between male and female content in this regard?* The connection between nudity and gender needs to be explored further. To look at this in more detail, it is possible to introduce interaction terms into the regression models. I use interactions between male/female and body exposure, but I also generate a gender ratio to interact with body exposure. The gender ratio is calculated as follows:

$$\text{Genderratio} = \frac{\text{no.female pics}}{\text{no.male pics} + \text{no.female pics}}$$

This gives a ratio varying from 0 to 1, with 1 being all female content and 0 being all male content. Including this into the analysis helps to measure the "degree" of male/female pictures and explore its effect. The results for marginal effects of the regression model are easier illustrated subsequently than the single gender estimations.

The results of Section 4.1 have shown that the average price per picture is higher for male accounts and the average revenue is higher for female accounts in the sample (due to higher posting frequency of branded content). Therefore, I introduce the logged values of price per picture to the regression model to shed light on potential differences in the effects on revenue and price. The dependent variable in Models 5 and 7 is the main dependent of this article lnAdrevenue (introduced in Section 3.2), and the results are compared with Models 6 and 8 and the dependent variable lnPicprice (see Table 9). In Models 5 and 6, the interactions between gender ratio and body exposure are estimated, and Models 7 and 8 contain a three-way interaction term of gender ratio and body exposure within different categories.

The main results refer to the interaction terms; the respective coefficients are difficult to directly interpret. Therefore, I follow the econometric standard and interpret the marginal effects. The marginal effects for the interaction between gender ratio and body exposure in Model 5 are displayed in Figure 4. All values are above zero and therefore significantly positive. That means that body exposure has a positive influence on advertising revenue for all contents, ranging from 0 (all contents male, left side) to 1 (all contents female, right side). The confidence intervals

TABLE 9 Hybrid models with interaction terms: Research question 2.

Variables	(5) lnAdrevenue (Genderratio interact)	(6) lnPicprice (Genderratio interact)	(7) lnAdrevenue (Genderratio category interact)	(8) lnPicprice (Genderratio category interact)
Mean(s) of kNumberofposts	0.046***	−0.057***	0.047***	−0.058***
	(0.011)	(0.013)	(0.011)	(0.013)
kNumberofposts (centered)	0.146***	0.002	0.146***	0.002
	(0.051)	(0.049)	(0.051)	(0.049)
Mean(s) of mFollowers	0.086***	0.067***	0.087***	0.067***
	(0.011)	(0.008)	(0.010)	(0.008)
mFollowers (centered)	0.017	0.092***	0.017	0.092***
	(0.031)	(0.027)	(0.031)	(0.027)
Mean(s) of Postengagement	0.227***	0.302***	0.225***	0.301***
	(0.019)	(0.017)	(0.019)	(0.017)
Post engagement % (centered)	0.229***	0.278***	0.229***	0.278***
	(0.021)	(0.017)	(0.021)	(0.017)
Body exposure # Genderratio	Figure 4	Figure 5		
Body exposure # Genderratio # category			Figure 6	Figure 7
HHI audience location	Incl.	Incl.	Incl.	Incl.
HHI audience language	Incl.	Incl.	Incl.	Incl.
HHI audience interests	Incl.	Incl.	Incl.	Incl.
Constant	7.547***	8.217***	7.278***	8.185***
	(0.285)	(0.193)	(0.360)	(0.279)
Observations	5110	6860	5110	6860
Number of id	435	446	435	446

Note: The dependent variable for Models 5 and 7 is the logged values for advertising revenue (main dependent). For Models 6 and 8, the logged values for the price per picture serve as dependent variable. In Models 5 and 7: $N = 5110$ and $n = 446$; only accounts with positive revenues and values for interaction terms are included within the estimation. In Models 6 and 8: $N = 6860$ and $n = 446$; whereas for some accounts, the information on post frequency is missing and the values for price per picture are available. Models 5 and 6 estimate the interaction between gender ratio and body exposure and Models 7 and 8 the interaction between gender ratio and body exposure per category. All models include independent variables for popularity and audience information based on Model 3 (Section 4.2.1).
All estimation are hybrid models with a between (between different accounts) and a cluster component (within the same account). *Mean(s)* are cluster-specific means, and *(centered)* are deviation scores between accounts to implement the hybrid model. Some variables are scaled in size (k in thousand, m in million). Robust standard errors are in parentheses.
***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

overlap indicating that there is no significant difference between male and female content in this regard. These results are contrasted by the marginal effects regarding the price per picture as dependent variable in Figure 5. The marginal effects are not significant regarding gender ratio 0.9 and 1 female content. This means that body exposure does not have significantly positive influence on the advertising price for completely female content.

The marginal effects show a significantly positive effect (all above zero) of body exposure on logged advertising revenue (y-axis) within the range from male (x-axis left, 0) to female (x-axis right, 1) content. Therefore, body

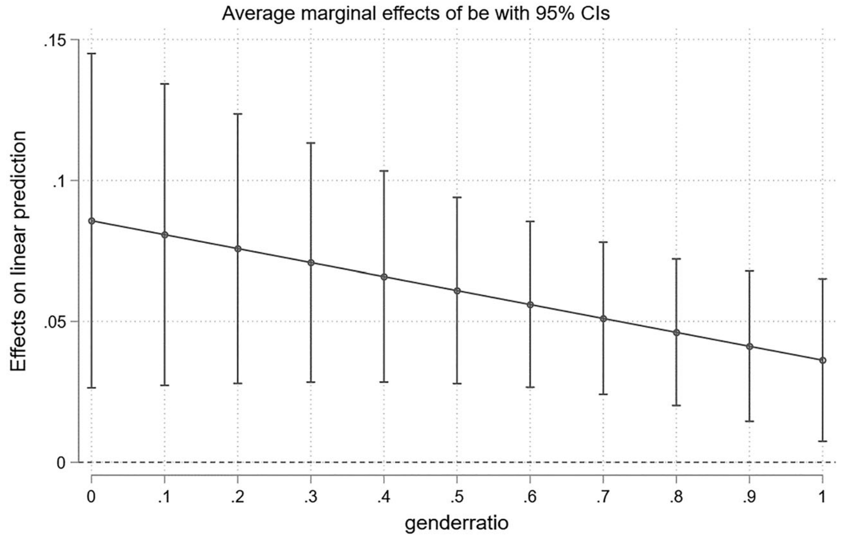

FIGURE 4 Marginal effects of Model 5 (Ad Revenue): Gender ratio and body exposure.

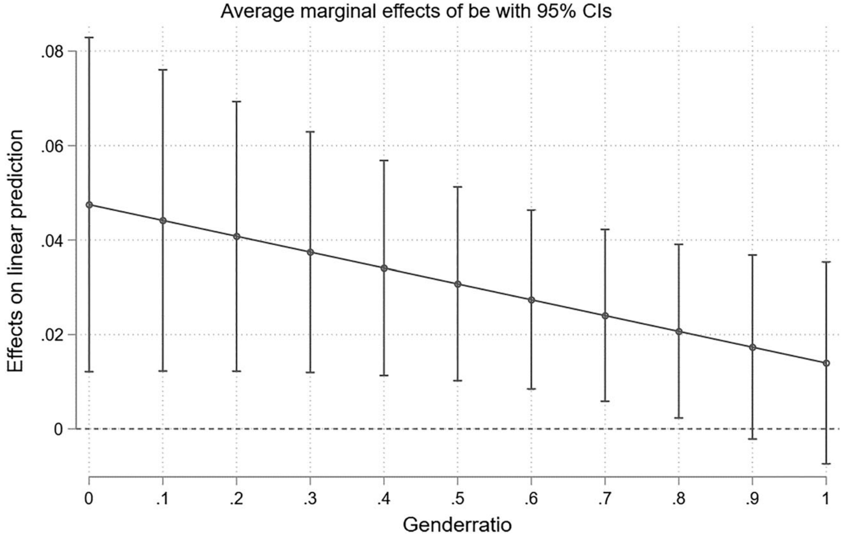

FIGURE 5 Marginal effects of Model 6 (Pic Price): Gender ratio and body exposure.

exposure has a significantly positive influence on lnAdrevene. There is no significant difference between male and female contents (confidence intervals overlap).

The marginal effects show a significantly positive effect (above zero) of body exposure on logged price per picture (y-axis) except for female content at the outer edge (90% and 100% female content, right). There is no significant difference between male and female contents (confidence intervals overlap). In difference to the dependent variable lnAdrevenue, body exposure is not significantly positive for female accounts regarding lnPicprice.

The marginal effects of Models 7 and 8 are illustrated in Figures 6 and 7. The three-way interaction term reduces the observations per gender in the respective category. Therefore, significant results for the overall sample

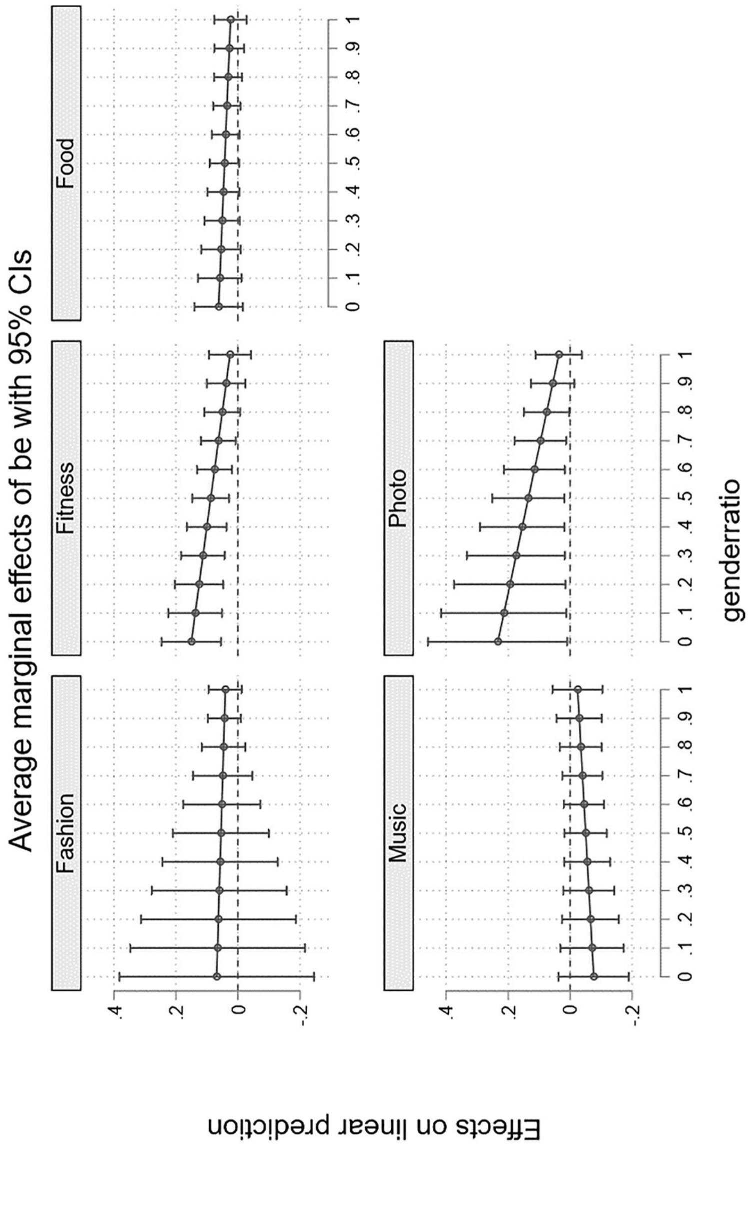

FIGURE 6 Marginal effects of Model 7 (Ad Revenue): Interaction per category.

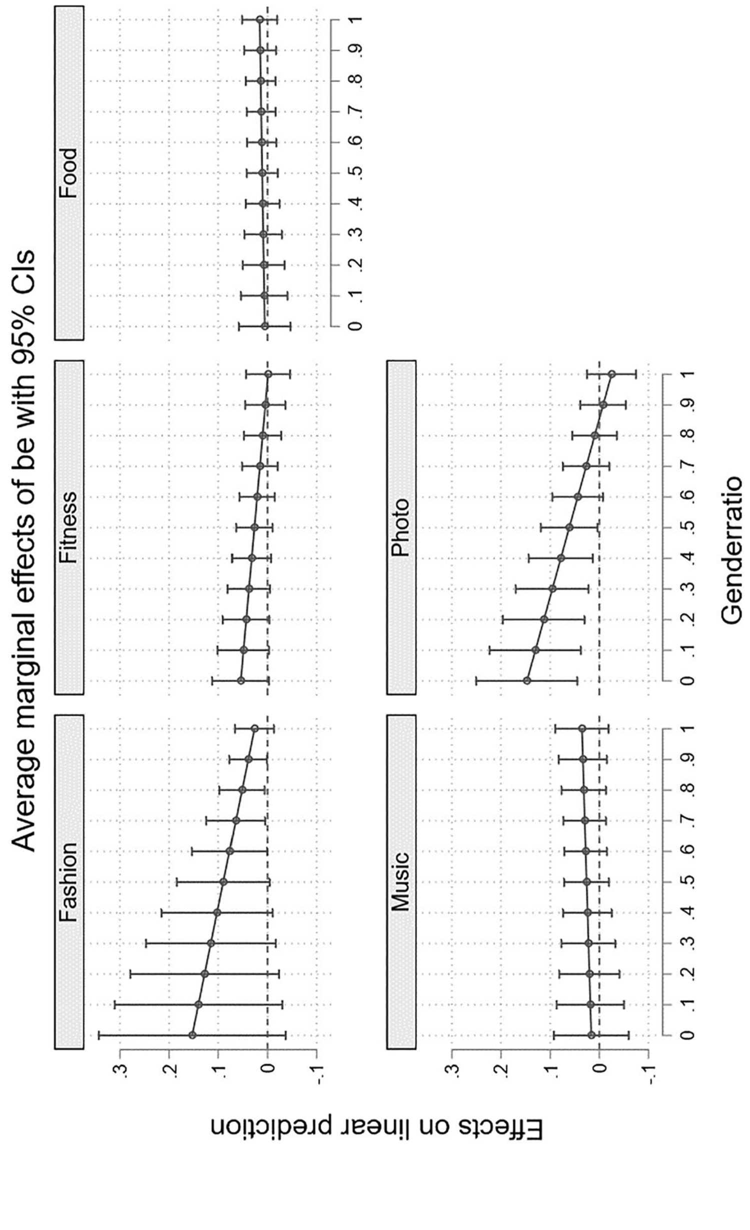

FIGURE 7 Marginal effects of Model 8 (Pic Price): Interaction per category.

(Figures 4 and 5) cannot be observed. It can be expected that the number of observations is too limited to observe systematic effects regarding the interaction between the three variables. However, the results show significantly positive effects for male content in the fitness category and slightly positive results for the category photo regarding the dependent variable lnAdRevenue. For accounts with male content, body exposure has a significantly positive effect on revenue. Regarding the dependent variable lnPicprice, only the category photo shows positive results for primarily male contents. Accounts with male content in this category are also positively influenced by body exposure. Confidence intervals between categories intersect and do not allow any interpretation of significance.

The marginal effects show only significantly positive effect within the category fitness (above zero) of body exposure on logged advertising revenue (y-axis). Therefore, body exposure has a significantly positive influence on lnAdrevene for male content in fitness. There is no significant difference between categories contents (confidence intervals overlap).

The marginal effects show no significantly positive effects (above zero) of body exposure on logged price per picture (y-axis) except for male in the category photo. There is no significant difference between categories contents (confidence intervals overlap).

4.3 | Robustness checks

4.3.1 | Hybrid versus fixed effects and random effects

When displaying a hybrid model, it is valuable to show the plain fixed effects and random effects for comparison and transparency. Also, the results can confirm estimations and robustness of results. That is why I present the results in Table 10. Model 9 shows the random effects model that can be compared with the fixed effects and the hybrid model (Models 10 and 11). Furthermore, I introduced a correlated random effects model (Mundlak, 1978). The model relaxes the assumption that there is no correlation between μ_i (the between error) and the time-variant variables ($\mu_i = \pi x_i + \nu_i$) (Schunck, 2013). Results are robust over all four models.

4.3.2 | Mediator variable

It can be argued that nudity does not directly influence the prices advertisers are willing to pay and thus advertising revenues, but the popularity among consumers and therefore (in a second step) prices revenues. Advertisers are mainly interested in the attention drawn towards their products and visibility among consumers. Being connected to nude or sexualized content can have positive but also negative effects on brand image (Wirtz et al., 2018). So, do advertisers choose to pay more for content with higher body exposure? The question can be answered using the popularity indicator as media variable between the independent and dependent variables. Figure 8 shows the construction of the mediator process.

The product method is the traditional approach for mediation analysis introduced by Baron and Kenny (Baron & Kenny, 1986; Bellavia, 2021). Three regressions are estimated to calculate an indirect effect of body exposure via popularity on prices.

1. $E[M|x,c] = \gamma_0 + \gamma_1 x + \gamma_2^T c$ *Body Exposure* affects *Popularity*
2. $E[Y|x,c] = \alpha_0 + \alpha_1 x + \alpha_2^T c$ *Popularity* affects *AdRevenue*
3. $E[Y|x,m,c] = \beta_0 + \beta_1 x + \beta_2 m + \alpha_2^T c$ *Body Exposure* affects *AdRevenue*

Indirect Effect $= \gamma_1 \beta_{2.}$

TABLE 10 Robustness checks: Random effects, fixed effects, hybrid, correlated random.

Variables	(9) lnAdRevenue random effects	(10) lnAdRevenue fixed effects	(11) lnAdRevenue hybrid	(12) lnAdRevenue correlated random
kNumberofposts	0.054***	0.147***		0.148***
	(0.013)	(0.052)		(0.051)
Mean(s) of kNumberofposts			0.052***	−0.096**
			(0.012)	(0.048)
kNumberofposts (centered)			0.148***	
			(0.051)	
mFollowers	0.080***	0.013		0.015
	(0.010)	(0.031)		(0.031)
Mean(s) of mFollowers			0.088***	0.073**
			(0.011)	(0.033)
mFollowers (centered)			0.015	
			(0.031)	
Post engagement %	0.231***	0.231***		0.231***
	(0.017)	(0.021)		(0.021)
Mean(s) of Postengagement			0.232***	0.001
			(0.020)	(0.028)
Post engagement % (centered)			0.231***	
			(0.021)	
Body exposure	0.039***	-	0.039***	0.039***
	(0.014)		(0.014)	(0.014)
Female pics	0.017	-	0.016	0.016
	(0.019)		(0.018)	(0.018)
Male pics	0.013	-	0.012	0.012
	(0.020)		(0.019)	(0.019)
Category	Incl.	-	Incl.	Incl.
HHI audience location	Incl.	Incl.	Incl.	Incl.
HHI audience language	Incl.	Incl.	Incl.	Incl.
HHI audience interests	Incl.	Incl.	Incl.	Incl.
Constant	7.639***	8.092***	7.566***	7.566***
	(0.308)	(0.258)	(0.311)	(0.311)
Observations	5320	5320	5320	5320
Overall *r*-squared	0.4258	0.1548	0.4265	0.4265
R-squared within	0.0897	0.0915	0.0915	0.0915
R-squared between	0.5055	0.1758	0.5057	0.5057
Number of id	454	454	454	454

Note: The dependent variable for all estimations in this table is *logged values of ad revenue*. These are robustness checks comparing random effects, fixed effects, hybrid, and correlated random models based on Model 3 (Section 4.2.1) $N = 5320$ and $n = 454$. *Mean(s)* are cluster-specific means, and *(centered)* are deviation scores between accounts to implement the hybrid model. Some variables are scaled in size (k in thousand, m in million). Robust standard errors are in parentheses.
***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

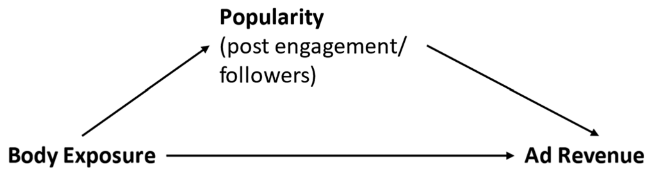

FIGURE 8 Mediator variable.

Regressing *Body Exposure* on *Followers* or *Postengagement* (3) failed to be significant in my study. Therefore, it can be argued that I do not observe a mediation effect, but that the effect of nudity directly and positively influences advertisers' choice and willingness to pay higher prices.

5 | CONCLUSION

5.1 | Synthesis and discussion of results

On the one hand, the *descriptive statistics* show that, on average, the prices for male content are higher than for female content within the sample. These results are interesting because they highlight complex dynamics of gender and possible inequality in the social media industry. This suggests a possible gender-based bias, with male content providers being able to command higher prices than female providers. It raises questions about the societal and cultural factors that may contribute to this bias, including gender stereotypes and expectations about the value of male versus female content. An in-depth analysis of gender effects, such as Marchenko and Sonnabend (2022) or Throsby et al. (2020) for artists, would be interesting to study effects in social media. A balanced data set regarding female and male content creators is needed to explore income (in)equalities in this dynamic and new digital industry. On the other hand, descriptive tests show that there are overall more female accounts in the sample and that these accounts post more branded content. Leading to the controversial result that, albeit the prices for male content are higher, accounts with female content perform comparably better regarding overall advertising revenue. This suggests that female creators may be more successful in monetizing their content or advertisers regarding them as a better fit for certain campaigns.

Beyond the descriptive results, the *analytic analysis* of the sample clearly shows a positive impact of body exposure on monetary success. Accounts with high level of body exposure achieve higher prices and advertising revenues than accounts with less nudity, regardless of the gender. From a social point of view, these results are interesting because of their potential reflection of societal values and norms in social media and beyond. It reflects body image and sexuality in modern media. The findings suggest that there is a considerable demand for sexualized content (also from advertising side), but it also raises questions about the objectification, body images, and perception of beauty. Moreover, the type of content and category must be taken into account. My results show that, in some categories, nudity is specifically successful, such as fitness. There might be a necessity to provide nude content within these fields to stay competitive as a content provider. The requirements for physical characteristics can differ in different categories.

Lastly, the effects of body exposure on (i) price per picture and (ii) ad revenues are compared. It is interesting to observe that body exposure has a significant positive impact on both aspects in general. However, comparing female to male content, body exposure does not positively influence the (i) price per picture for "all female" (90% or 100% female content). This can be due to the fact that these accounts in comparison already provide large amounts of body exposure in the sample (as the descriptive results suggest). The stars may already offer so much nudity that there is no marginal effect of body exposure. Unfortunately, the data set does not provide enough information to observe marginal effects within an account, but only a comparison between accounts. A more extensive data set including dynamic information on nudity could reveal interesting insights on the optimal level of nudity on an Instagram account.

Overall, the clear positive effects of body exposure on monetary success on Instagram suggest that there is a need for greater awareness and dialogue around the social and cultural implications of sexualized content. These findings provide first insights for a fact-based economic and societal discussion in a complex and multifaceted topic regarding societal values and norms, as well as ethical and legal standards. Instagram stars may use the opportunity of nude content to generate higher social media income. At the same time, the sustainability of the economic model is questionable because of concerns regarding ethical and social implications of sexualized content. Especially the impact of this content on children and adolescents who are increasingly exposed to it through social media should be considered.

5.2 | Concluding remarks and limitations

This paper empirically studies the effects of nudity on Instagram success. Stars need to invest heavily into audience building, by drawing attention to themselves (audience attraction) and managing attention in the long run (audience maintenance). Body exposure is studied as part of audience attraction within this paper. To answer the research question *Does body exposure drive income success on Instagram?* it can be said that body exposure positively influences advertising income of Instagram creators in the sample. A descriptive study of the data shows that a high level of nudity is not exclusively necessary for top income positions but is still common among the big ones. A more detailed statistical analysis shows that accounts with high level of body exposure achieve higher prices and advertising revenues in comparison. In summary, "sex sells" as one more nude picture generates on average 4% higher income when comparing body exposure between accounts.

The second question was: *Is there a difference between male and female content in this regard?* On the one hand, some differences can be observed, for instance, a *T*-test shows that body exposure is significantly higher for accounts with focus on female content. Regarding monetary success, male content achieves higher prices, whereas female accounts provide more branded content and eventually achieve higher advertising revenues. On the other hand, the clear trend is "nudity has positive impact no matter what gender"; both female and male accounts profit from body exposure. In this sample, women provide more nude content and post more branded pictures. If nudity is a type of investment/cost from an economic perspective, do women need to invest more to achieve high levels of revenue?

The results can only be interpreted carefully and used as a first indicator; there is no clear evidence on gender differences, or a gender pay gap. More research is needed to study differences in income, between gender, categories, platforms, and so forth. There are some limitations to this study, which open possibilities for further studies. As mentioned in the beginning, there are no traditional gatekeepers in social media markets (although agencies like multi-channel networks established themselves during the last years (Budzinski & Gaenssle, 2020)). Yet, there are new gatekeepers, first and foremost the platform immanent algorithms. Sophisticated recommendation services manage scarce consumer attention and pre-select content according to individual preferences (Budzinski et al., 2021). A study by Richard et al. (2020) shows indicators that Instagram biases content allocation and favors nude content. According to their results, Instagram prioritizes contents with high body exposure over others and pushes those posts. Because the algorithms are business secrets and ever-changing, it is not possible to subtract its influence on income in my study. However, if this is the case, the gap between male and female income is expected to be even bigger—with females' higher degree in body exposure, further algorithmic support of the platform and, yet, no difference in payment. In any case, algorithm management and upload behavior matter, as already pointed out by Budzinski and Gaenssle (2020) and Gaenssle and Budzinski (2021).

Nonetheless, it must be considered that this study only focuses on the top group in five categories, which are dominated by female content (51.9% female content; 21.6% male content). A specific sampling according to gender would be interesting, to compare representative groups. From an economic point of view, scarcity for male content is higher than for female content, which is why it can be argued that it is more costly to advertise with them.

There are multiple sources of income for social media stars (see Section 3.2). In this paper, only the advertising revenue is observable. It serves as a good proxy, because it is common practice to earn money from ads and brand communication on Instagram. However, especially for nudity, there might be other sources of income. The website Onlyfans became very popular over the last years, featuring predominantly adult entertainment content. For very popular accounts with high BE, this might be a source of income, with Instagram merely being a marketing platform. Popular accounts link to Onlyfans, where a payment barrier is established to access further contents. Although the income proxy of this paper is legitimate, further studies on other sources of income would be interesting (if data are available).

There are limits to the data used in this paper. I use estimations of prices; the exact numbers are not available. The data seem very reliable, conforming effects of past studies (see Section 4.2). Lastly, there is no classification of sex appeals within this data set. I do not differentiate, for example, between sportswear and lingerie. The sexual appeal of the pictures varies. Yet, sexual appeals are very subjective and leave room for further biases in coding.

ACKNOWLEDGMENTS

I thank the *Department of Economic Science and Media* of the Ilmenau University of Technology for their support and young researcher funding. I thank Oliver Budzinski and Philipp Kunz-Kaltenhäuser for valuable comments on an earlier version of this paper and Rohangis Mohseni for helpful suggestions regarding methods. Moreover, thanks to all participants commenting on this paper at the ACEI conference 2021, Weizenbaum Institut “Plamadiso Talk” 2021, European Media Management Association Annual Conference 2022, Digital Europe Economic Seminars 2022.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Sophia Gaenssle https://orcid.org/0000-0003-1283-3862

REFERENCES

Adler, M. (1985). Stardom and talent. *The American Economic Review*, *75*(1), 208–212.

Adler, M. (2006). Stardom and talent. In V. A. Ginsburgh & D. Throsby (Eds.), *Handbook of the economics of art and culture* (Vol. 1) (pp. 895–906). Elsevier. https://doi.org/10.1016/S1574-0676(06)01025-8

Allison, P. D. (2009). Fixed effects regression models. In *Quantitative applications in the social sciences* (Vol. 160). SAGE. https://doi.org/Paul

Armstrong, M. (2006). Competition in two-sided markets. *The Rand Journal of Economics*, *37*(3), 668–691. https://doi.org/10.1111/j.1756-2171.2006.tb00037.x

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037//0022-3514.51.6.1173

Bellavia, A. (2021). *Mediation and interaction analysis: Introduction and overview of Stata commands*. https://www.stata.com/symposiums/biostatistics-and-epidemiology21/slides/Bio21_Bellavia.pdf

Benoit, K. (2011). *Linear regression models with logarithmic transformations*. London School of Economics.

Bryson, A., Rossi, G., & Simmons, R. (2014). The migrant wage premium in professional football: A superstar effect? *Kyklos*, *67*(1), 12–28. https://doi.org/10.1111/kykl.12041

Budzinski, O., & Gaenssle, S. (2020). The economics of social media (super-)stars: An empirical investigation of stardom and success on YouTube. *Journal of Media Economics*, *31*(3–4), 75–95. https://doi.org/10.1080/08997764.2020.1849228

Budzinski, O., Gaenssle, S., & Lindstädt-Dreusicke, N. (2021). Data (R)Evolution—The Economics of algorithmic search and recommender services. In S. Baumann (Ed.), *Handbook of digital business ecosystems (forthcoming)*. Elgar. https://doi.org/10.2139/ssrn.3769386

Budzinski, O., & Pannicke, J. (2017). Does popularity matter in a TV song competition? Evidence from a National Music Contest. *Ilmenau Discussion Papers*, *21*(106). https://doi.org/10.2139/ssrn.2964240

Burgess, J., & Green, J. (2018). YouTube: Online video and participatory culture. In *Digital media and society* (2nd ed.). Polity Press.

Candela, G., Castellani, M., Pattitoni, P., & Di Lascio, F. M. L. (2016). On Rosen's and Adler's hypotheses in the modern and contemporary visual art market. *Empirical Economics*, *51*(1), 415–437. https://doi.org/10.1007/s00181-015-1002-3

Che, Y.-K., & Mierendorff, K. (2019). Optimal dynamic allocation of attention. *American Economic Review*, *109*(8), 2993–3029. https://doi.org/10.1257/aer.20171000

Crain, W. M., & Tollison, R. D. (2002). Consumer choice and the popular music industry: A test of the superstar theory. *Empirica*, *29*(1), 1–9. https://doi.org/10.1023/A:1014651130414

Dudley, S. C. (1999). Consumer attitudes toward nudity in advertising. *Journal of Marketing Theory and Practice*, *7*(4), 89–96. https://doi.org/10.1080/10696679.1999.11501854

Ehrmann, T., Meiseberg, B., & Ritz, C. (2009). Superstar effects in deluxe gastronomy—An empirical analysis of value creation in German quality restaurants. *Kyklos*, *62*(4), 526–541. https://doi.org/10.1111/J.1467-6435.2009.00449.X

Evans, D. S., & Schmalensee, R. (2007). Industrial organization of markets with two-sided platforms. *Competition Policy International*, *3*(1), 151–179.

Facebook. (2021). *Instagram community guidelines*. Facebook. https://www.facebook.com/help/instagram/477434105621119/

Filimon, N., López-Sintas, J., & Padrós-Reig, C. (2011). A test of Rosen's and Adler's theories of superstars. *Journal of Cultural Economics*, *35*(2), 137–161. https://doi.org/10.1007/s10824-010-9135-x

Franck, G. (2019). The economy of attention. *Journal of Sociology*, *55*(1), 8–19. https://doi.org/10.1177/1440783318811778

Franck, E., & Nüesch, S. (2007). Avoiding 'star wars'—Celebrity creation as media strategy. *Kyklos*, *60*(2), 211–230. https://doi.org/10.1111/j.1467-6435.2007.00369.x

Franck, E., & Nüesch, S. (2008). Mechanisms of superstar formation in German soccer: Empirical evidence. *European Sport Management Quarterly*, *8*(2), 145–164. https://doi.org/10.1080/16184740802024450

Franck, E., & Nüesch, S. (2012). Talent and/or popularity: What does it take to be a superstar? *Economic Inquiry*, *50*(1), 202–216. https://doi.org/10.1111/j.1465-7295.2010.00360.x

Gaenssle, S., & Budzinski, O. (2021). Stars in social media: New light through old windows? *Journal of Media Business Studies*, *18*(2), 79–105. https://doi.org/10.1080/16522354.2020.1738694

Giles, D. E. (2006). Superstardom in the US popular music industry revisited. *Economics Letters*, *92*(1), 68–74. https://doi.org/10.1016/j.econlet.2006.01.022

Gräve, J.-F. (2017). Exploring the perception of influencers vs. traditional celebrities. In Association for Computing Machinery. (Ed.), *Proceedings of the 8th international conference on social media & society—#SMSociety17* (pp. 1–5). ACM Press. https://doi.org/10.1145/3097286.3097322

Hofmann, K. H., & Opitz, C. (2019). Talent and publicity as determinants of superstar incomes: Empirical evidence from the motion picture industry. *Applied Economics*, *51*(13), 1383–1395. https://doi.org/10.1080/00036846.2018.1527452

Hypeauditor. (2020). *The state of influencer marketing 2020: Influencer fraud dynamics*. https://hypeauditor.com/blog/state-2020/

Jung, C., & Nüesch, S. (2019). The more others care, the more you share?—Social contagion as a stardom trigger of social media superstars. *Applied Economics*, *51*(9), 881–888. https://doi.org/10.1080/00036846.2018.1497849

Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. *The American Economic Review*, *75*(3), 424–440. http://www.jstor.org/stable/1814809

Katz, M. L., & Shapiro, C. (1994). Systems competition and network effects. *Journal of Economic Perspectives*, *8*(2), 93–115. https://doi.org/10.1257/jep.8.2.93

Krueger, A. B. (2005). The economics of real superstars: The market for rock concerts in the material world. *Journal of Labor Economics*, *23*(1), 1–30. https://doi.org/10.1086/425431

Lehmann, E. E., & Schulze, G. G. (2008). What does it take to be a star?—The role of performance and the media for German soccer players. *Applied Economics Quarterly*, *54*(1), 59–70. https://doi.org/10.3790/aeq.54.1.59

Leibenstein, H. (1950). Bandwagon, snob, and Veblen effects in the theory of consumers' demand. *Quarterly Journal of Economics*, *64*(2), 183–207. https://doi.org/10.2307/1882692

Lorber, J. (1996). Beyond the binaries: Depolarizing the categories of sex, sexuality, and gender. *Sociological Inquiry*, *66*(2), 143–160. https://doi.org/10.1111/j.1475-682X.1996.tb00214.x

Lorber, J. (2005). *Breaking the bowls: Degendering and feminist change*. WW Norton New York.

Lucifora, C., & Simmons, R. (2003). Superstar effects in sport: Evidence from Italian soccer. *Journal of Sports Economics*, *4*(1), 35–55. https://doi.org/10.1177/1527002502239657

Lull, R. B., & Bushman, B. J. (2015). Do sex and violence sell? A meta-analytic review of the effects of sexual and violent media and ad content on memory, attitudes, and buying intentions. *Psychological Bulletin*, *141*(5), 1022–1048. https://doi.org/10.1037/bul0000018

Marchenko, M., & Sonnabend, H. (2022). Artists' labour market and gender: Evidence from German visual artists. *Kyklos*, *75*, 456–471. https://doi.org/10.1111/kykl.12302

Marwick, A. E. (2016). You may know me from YouTube: (Micro-)celebrity in social media. In P. D. Marshall & S. Redmond (Eds.), *A companion to celebrity* (pp. 333–350). Wiley Blackwell.

Meiseberg, B. (2014). Trust the artist versus trust the tale: Performance implications of talent and self-marketing in folk music. *Journal of Cultural Economics*, *38*(1), 9–42. https://doi.org/10.1007/s10824-012-9196-0

Mittal, B., & Lassar, W. M. (2000). Sexual liberalism as a determinant of consumer response to sex in advertising. *Journal of Business and Psychology*, *15*(1), 111–127. https://doi.org/10.1023/A:1007723003376

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, *46*(1), 69. https://doi.org/10.2307/1913646

Neuhaus, J. M., & Kalbfleisch, J. D. (1998). Between- and within-cluster covariate effects in the analysis of clustered data. *Biometrics*, *54*(2), 638. https://doi.org/10.2307/3109770

Reichert, T. (2002). Sex in advertising research: A review of content, effects, and functions of sexual information in consumer advertising. *Annual Review of Sex Research*, *13*(1), 241–273. https://doi.org/10.1080/10532528.2002.10559806

Richard, É., Duportail, J., Kayser-Bril, N., & Schacht, K. (2020). *Undress or fail: Instagram's algorithm strong-arms users into showing skin*. https://algorithmwatch.org/en/instagram-algorithm-nudity/

Rochet, J., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, *1*(4), 990–1029. https://doi.org/10.1162/154247603322493212

Rochet, J., & Tirole, J. (2006). Two-sided markets: A progress report. *The Rand Journal of Economics*, *37*(3), 645–667. https://doi.org/10.1111/j.1756-2171.2006.tb00036.x

Rosen, S. (1981). The economics of superstars. *American Economic Review*, *71*(5), 845–858.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*(5762), 854–856. https://doi.org/10.1126/science.1121066

Schunck, R. (2013). Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. *The Stata Journal: Promoting Communications on Statistics and Stata*, *13*(1), 65–76. https://doi.org/10.1177/1536867X1301300105

Severn, J., Belch, G. E., & Belch, M. A. (1990). The effects of sexual and non-sexual advertising appeals and information level on cognitive processing and communication effectiveness. *Journal of Advertising*, *19*(1), 14–22. https://doi.org/10.1080/00913367.1990.10673176

Shapiro, C., & Varian, H. R. (1999). *Information rules: A strategic guide to the network economy*. Harvard Business School Press.

Simon, H. A. (1971). Designing organizations for an information rich world. In M. Greenberger (Ed.), *Computers, communications, and the public interest* (pp. 37–72). Johns Hopkins P.

Stigler, G. J., & Becker, G. S. (1977). De Gustibus Non Est Disputandum. *American Economic Review*, *67*(2), 76–90.

Taylor & Greg. (2014). Scarcity of attention for a medium of abundance. In M. Graham & W. H. Dutton (Eds.), *Society and the internet: How networks of information and communication are changing our lives* (pp. 257–271). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199661992.003.0017

Throsby, D., Petetskaya, K., & Shin, S. Y. (2020). *The gender pay gap among Australian artists: Some preliminary findings*. Australia Council for the Arts.

Torgler, B., Antic, N., & Dulleck, U. (2008). Mirror, mirror on the wall, who is the happiest of them all? *Kyklos*, *61*, 309–319. https://doi.org/10.1111/j.1467-6435.2008.00403.x

Wirtz, J. G., Sparks, J. V., & Zimbres, T. M. (2018). The effect of exposure to sexual appeals in advertisements on memory, attitude, and purchase intention: A meta-analytic review. *International Journal of Advertising*, *37*(2), 168–198. https://doi.org/10.1080/02650487.2017.1334996

How to cite this article: Gaenssle, S. (2023). Income distribution and nudity on social media: Attention economics of Instagram stars. *Kyklos*, 1–29. https://doi.org/10.1111/kykl.12363

APPENDIX A

Figure A1 is coded as follows: Coding BE: 0–12

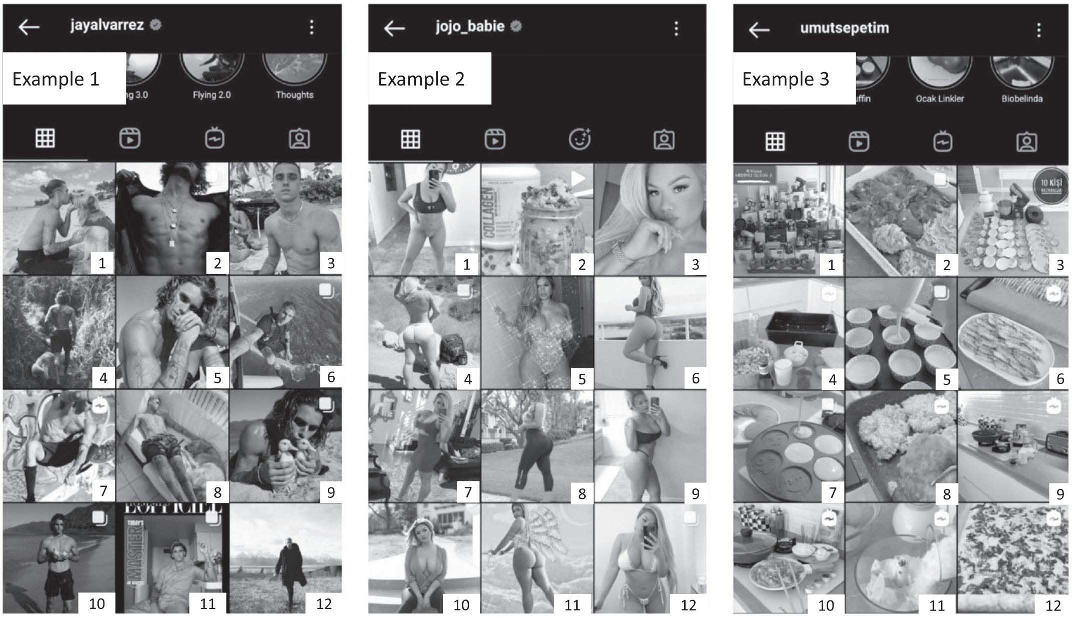

FIGURE A1 Section 3.3 Example for coding BE.

Coding Gender: 0–12, in (female, male, mixed, ambivalent, no identification)

Example 1:
BE: 8 *(1 = 1; 2 = 1; 3 = 1; 4 = 1; 5 = 1; 6 = 0; 7 = 1; 8 = 1; 9 = 0; 10 = 1; 11 = 0; 12 = 0)*
Gender: 0-12-0-0-0 *(all male)*

Example 2:
BE: 10 *(1 = 1; 2 = 0; 3 = 0; 4 = 1; 5 = 1; 6 = 1; 7 = 1; 8 = 1; 9 = 1; 10 = 1; 11 = 1; 12 = 1)* with pictures 7 and 8 dressed, but 50% or more focus is on primary sex characteristics
Gender: 11-0-0-0-1 *(all female, except second picture "no identification")*

Example 3:
BE: 0 *(no nudity in any picture)*
Gender: 0-0-0-0-12 *(all "no identification," no human characteristics)*

TABLE A1 Section 4.2.1: Correlation popularity factors.

Variables	(1)	(2)	(3)
(1) Likes	1.000		
(2) Followers	0.570	1.000	
(3) Postengagement	0.603	−0.016	1.000