A Matter of Taste: The Negative Welfare Effect of Expert Judgments*

Nicolas Lagios † Pierre-Guillaume Méon‡
Job Market Paper

January 23, 2024
[Click here for the latest version]

Abstract

Expert judgments may increase or decrease consumer welfare depending on experts' ability to redirect consumers toward goods they enjoy. Leveraging the discontinuity created by the attribution of the Booker Prize, a leading literary award, we confirm that the prize attracts readers to consumption. We then investigate how it affects consumer surplus. We measure consumer *ex post* satisfaction from reading a book by the sentiment and the rating of the reviews posted on Amazon. We show that the Booker reduces satisfaction and that this negative effect is driven by a misalignment between the tastes of the jury and those of consumers. We quantify the associated loss in welfare by calibrating a structural model of demand. We find that the prize reduces consumer surplus by USD135,000 annually, meaning that a consumer buying a Booker Prize-winning book experiences a loss in surplus of 4% of the average price of a book.

Keywords: Awards, Prizes, Welfare, Sales, Experts, Books, Consumer Surplus.
JEL Classifications: D12, D83, L15, L82, Z11.

*We would like to thank Constantin Lagios, Virginie Bruneau, Sandra Rothenberger, and the participants of the 2022 Beyond Basic Questions workshop in Bern, the 2022 European Workshop on Applied Cultural Economics in Turin, the 2022 Edition of the Sapienza Economics Ph.D. Summer Meeting in Rome, the 2023 Meeting of the European Public Choice Society in Hanover, the 2023 Conference on Economic Design in Girona, the 2023 journées Louis-André Gérard-Varet in Marseille, the 2023 meeting of the European Economic Association in Barcelona, and the 2023 Back-to-School event of DULBEA in Brussels for useful comments. Any errors are our own.

†Université libre de Bruxelles (ULB), CEBRIG, DULBEA, Research Fellow F.R.S.-FNRS - Aspirant FNRS, CP-114/03, avenue F.D. Roosevelt 50, 1050 Bruxelles, Belgium. (nicolas.lagios@ulb.be).

‡Université libre de Bruxelles (ULB), CEBRIG, DULBEA, CP-114/03, avenue F.D. Roosevelt 50, 1050 Bruxelles, Belgium. (p-guillaume.meon@ulb.be).

1 Introduction

From the glitz and glamour of film festivals to the sophistication of wine or culinary awards, expert judgments can drive consumers to or away from the products they review (Ginsburgh [2003], Ashenfelter and Jones [2013], English [2014]). Those judgments are particularly important for experience goods, the utility of which consumers, by definition, cannot know prior to consumption. By assessing those goods and sharing their judgments with the public, experts can send a quality signal allowing consumers to choose better goods, thereby delivering welfare gains.

However, that view of the work of experts as welfare-enhancing rests on the assumption that their judgments reflect the tastes of consumers or, to put it simply, that they can tell consumers what they will like. This assumption is questionable on several grounds. First, telling others what they will like supposes an interpersonal comparison of likes and dislikes, against which both economics and psychology warn. Robbins [1938, p. 637], citing Jevons, reminds us, "Every mind is inscrutable to every other mind and no common denominator of feeling is possible." Bartoshuk [2014] makes the same point and emphasizes that psychological research shows systematic differences across individuals in the perception of pleasure.

Second, expert taste may differ from those of laypeople in a systematic way, as the sociological analysis of Bourdieu [1979, 1983] suggests. He argues that "experts have specific dispositions (habitus) shaped by their social trajectory" (Bourdieu [1983, p. 311]). In other words, people's tastes are not exogenously given but determined by their personal history and their position in society. Insofar as the personal history and the position in society of experts differ from those of laypeople, their tastes are likely different, too. What is more, the objective of experts may not be so much to put their approval on the goods that laypeople will like but to establish or foster their own legitimacy in their field (Bourdieu [1983]), which may give them an incentive to support products that are at odds with the tastes of most of the public. The political economy of experts further suggests that signaling what the public will like may not be their main objective. Experts are closer to the industries that they assess than laypeople, which can influence their judgment (Dobrescu *et al.* [2013]). Firms, advertisers, or commercial interest can try to capture experts' attention and praise (Cameron [1995]). Members of the juries awarding prizes are notoriously courted or lobbied by filmmakers, writers, or publishers (English [2014]).

As a result, the alignment of the tastes of experts and laypeople cannot be taken for granted, especially for goods that cannot be objectively assessed, like wine and art (Ginsburgh [2016]). In line with that presumption, there is evidence of systematic discrepancies between the judgment of experts and that of laypeople about classical music (Asmat *et al.* [2023]), popular music (Haan *et al.* [2005]), movies (Holbrook [1999]), architecture (Coeterier, 2002), landscape (Rogge *et al.* [2007]), and books (Lagios and Méon [2023]).

If the tastes of experts are indeed poorly aligned with those of consumers, the latter may end up consuming products they do not like and that they would not have initially consumed, hence leading consumers to experience lower utility. Contrary to conventional wisdom, expert judgments may, therefore, decrease consumer welfare. The aim of this paper is to determine whether this is the case and, if so, to quantify the resulting welfare loss.

To address that question, we focus on the Booker Prize, an internationally known literary prize awarded annually since 1969 by a committee of literary experts. Literary prizes provide an ideal case to study the welfare effects of expert judgments for at least two reasons. First, literary prizes are one of the main sources of expert judgments in the book industry and are central to the production and reception of books (English [2014]). Second, the book industry is characterized by a wide range of choices which, combined to the experience good nature of books, makes purchasing decisions complex and hazardous for consumers. This means that pre-purchase information, such as prizes, affects consumption choices (Ponzo and Scoppa [2015], Lagios and Méon [2023]).

We begin our analysis by investigating how the Booker affects the demand for books, as the prize will affect consumer surplus only if it influences consumption choices. To that end, we construct a rich dataset that covers the near universe of books published between 2015 to 2021, which is the period over which we can track the entire daily sales of a book on Amazon. Our dataset includes daily information on Amazon sales ranks, prices, and ratings for nearly 171,000 editions coming from more than 58,000 books. As Amazon's market share in physical and electronic books in the US is 42% and 89%, respectively, our data capture a significant part of the US book market.[1]

When estimating the effect of the Booker on sales, the main challenge is that a book's unobservable characteristics may drive both the probability of winning an award and commercial success. The jury may, for example, deliberately pick books that will likely be successful or incidentally reward characteristics that make a book successful. A

[1] See https://www.t4.ai/companies/amazon-market-share and https://wordsrated.com/book-sales-statistics.

naive regression of sales ranks on the Booker is therefore unlikely to reflect a causal effect.

We address endogeneity by elaborating on the discontinuity-based strategy used by Reimers and Waldfogel [2021] in their study on the effect of reviews. Specifically, we examine whether the discontinuity over time in public attention on a book prompted by the attribution of the Booker leads to a discontinuity in daily sales for that book. We find that it does and that the impact is substantial. In particular, the Booker raises annual book sales by 50%, on average, which corresponds to 216,000 additional copies.

We then leverage the customer reviews posted on Amazon to investigate the causal impact of the Booker on consumer satisfaction, as the impact of the prize on consumer surplus depends on its ability to redirect consumers toward books they will enjoy. Our sample consists of 6.7 million reviews. We measure consumer *ex post* satisfaction from reading a book in two ways. First, we perform a sentiment analysis on the textual content of each review. Sentiment analysis is a natural language processing technique that extracts the sentiment valence of an opinionated text, which can range from negative to positive (Pang and Lee [2008]), thereby providing a measure of satisfaction. We then confirm the results obtained with the sentiment analysis by using the review rating (number of stars), whereby a higher rating indicates a higher consumer *ex post* satisfaction.

We gauge the effect of the Booker on consumer satisfaction using a difference-in-differences design, where we compare how sentiments and ratings for awarded and non-awarded books change after the attribution of the prize. In line with the presumption that experts may redirect consumers to products that they do not enjoy, we observe that the Booker increases the probability of a book receiving a negative review and decreases its rating. Accordingly, the Booker negatively affects consumer *ex post* satisfaction and, hence, surplus. These findings stand up to a series of robustness checks, including using alternative econometric approaches such as regression discontinuity design and instrumental variable.

We then report a series of findings suggesting that the negative effect of the Booker on consumer satisfaction is indeed a matter of taste, specifically that it is driven by a misalignment in taste between the members of the jury and the public. We first replicate our baseline analyses, but this time we focus on a prize awarded by readers: the Goodreads Choice Award for Fiction. Like the Booker, the Goodreads prize provides visibility to a book and is a signal of quality. Unlike the Booker, however, the prize is awarded by a jury of laypeople whose tastes are arguably closer to those of the average reader, which should therefore result in less dissatisfaction. Supporting this idea, our results show that the Goodreads prize has no negative effect on satisfaction despite boosting sales.

We then leverage the variations in the Booker jury across editions to assess how its composition affects consumer satisfaction. Specifically, as many jury members are authors themselves, we can condition the effect of the prize on the rating given by readers to the books written by those jury members. The idea behind this approach is that if judges are able to write books that appeal to consumers, then they might be more likely to select a book that consumers will also like. In other terms, we use the rating of the books written by the jury as a proxy for their ability to award a book consumers will enjoy, either because they have the same tastes or because they can correctly predict them. Our findings confirm this premise: When a given year's judges' books receive higher ratings by readers, the effect of the prize on satisfaction is less negative and even becomes indistinguishable from zero for very high ratings. Furthermore, we show that when the cultural proximity of the jury members with the readers is higher – which can be interpreted as implying closer tastes with readers – the effect of the prize is also less negative. In a third series of tests, we use an online survey to document that respondents often report to be disappointed in awarded books and that many of them blame their discontent on a misalignment of the tastes of jury members with theirs. Overall, this series of findings suggest that the negative effect of the Booker on satisfaction is driven by the distance between the tastes of the jury members and those of readers. This is consistent with a model where prizes, regardless of the composition of their jury, attract readers to consumption, but the latter may be disappointed if their tastes are too far from those of the jury.

Last but not least, we quantify the loss in welfare induced by the Booker. To do so, we calibrate a structural model of demand for books in which the surplus of consumers depends on the difference between their expectations regarding the utility a book will give them ("decision utility") and the true utility they get from it ("experienced utility"; Kahneman [1994], Allcott [2011, 2013]). Specifically, our welfare analysis rests on the comparison of consumer surplus under two scenarios: a status quo scenario where consumers can use the prize to gauge the book and a counterfactual scenario in which the prize does not exist. Our lower bound and most conservative estimates, which assume that absent the Booker consumers correctly assess a book's utility, suggest that the prize reduces consumer surplus by USD135,000 each year. This means that a consumer buying a Booker Prize-winning book experiences a loss in her surplus of

USD0.60, which is non-negligeable as it corresponds to 4 percent of the average price of a book. We further show that this loss in welfare mainly arises from consumers switching from non-awarded to awarded books that they expect to enjoy more – that is, a business-stealing effect – rather than from consumers expanding their total book consumption.

This paper contributes to several strands of literature. The first is the literature on awards and prizes, to which we contribute by confirming that awards increase commercial success (Ashworth *et al.* [2010], Ponzo and Scoppa [2015], Ginsburgh *et al.* [2019], Lagios and Méon [2023]). We also provide additional evidence on a more recent finding of that literature, which is that awards can deteriorate online reviews posted by users (Rossi [2021], Lagios and Méon [2023]), thereby also contributing to the burgeoning literature on online reviews and rating systems (Hörner and Lambert [2021], Reimers and Waldfogel [2021], Acemoglu *et al.* [2022]).

The main contribution of the paper to these strands of literature, however, is twofold. First, we provide evidence that the negative effect of awards on users' reviews is driven by a misalignment of the tastes of experts with those of consumers. Second, by leveraging the distinction between "decisions utility" and "experienced utility", we show that this divergence of tastes results in a welfare loss for consumers, which we subsequently quantify. In that respect, the paper more generally contributes to the literature on experts (Ginsburgh and van Ours [2003], Reinstein and Snyder [2005], Hilger *et al.* [2011], Friberg and Grönqvist [2012], Loeper *et al.* [2014], Ginsburgh *et al.* [2019], Reimers and Waldfogel [2021]) by showing how their judgments can affect commercial success and consumer welfare. An important implication of our analysis is to flesh out the view of experts as the agents of consumers who are the principals (Cameron [1995]) and recall that the principal-agent relationship can be suboptimal if their interests are not aligned.

More generally, our findings qualify the notion of quality when applied to experience goods. Previous research has proxied quality by sales (Deuchert *et al.* [2005]) or best-of lists (Ginsburgh [2003], Ginsburgh and Weyers [2014]). Our findings underline, by contrast, that quality can only be assessed with respect to a given set of preferences and tastes. Moreover, our findings show that commercial success does not guarantee quality, as defined as the capacity to maximize consumers utility, because goods that are imperfectly aligned with the tastes of consumers can nonetheless be commercially successful.

The remainder of the paper is organized as follows. Section 2 describes our theoretical framework. Section 3 provides background information on the Booker and its functioning. Section 4 presents the data sources and detailed descriptives. Section 5 investigates the effect of the Booker on sales, while Section 6 explores its impact on consumer satisfaction. Section 7 reports evidence that the negative effect of the prize on satisfaction is driven by a misalignment of the tastes of the jury with those of readers. Section 8 calibrates a structural demand model to quantify the loss in consumer surplus induced by the Booker. Section 9 concludes.

2 Theory: How Prizes Can Affect Consumer Welfare

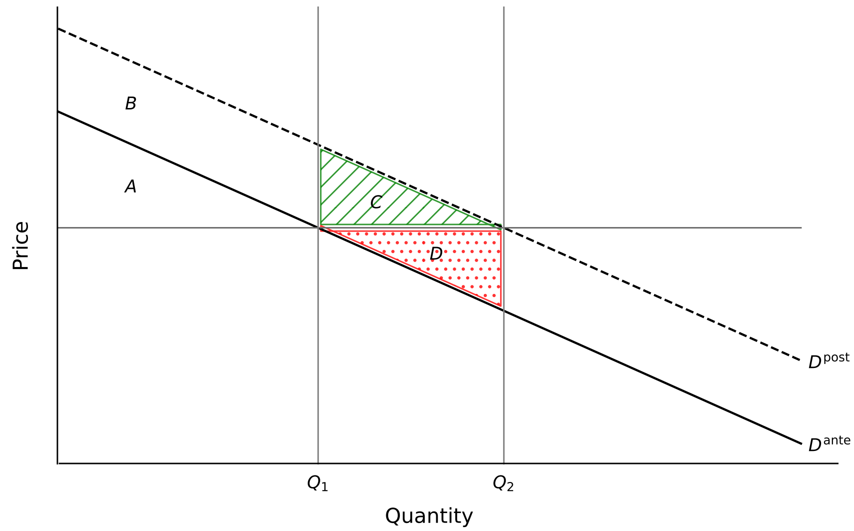
To describe the potential welfare effects of literary prizes, we follow Jin and Sorensen [2006], Allcott [2011], Train [2015], and Reimers and Waldfogel [2021] and distinguish between *ex ante* expected utility, or decision utility following Kahneman [1994], and *ex post* experienced utility.

Because books are experience goods, consumers are *ex ante* imperfectly informed of the utility they will get from a book that they are planning to purchase (Nelson [1970]). They therefore form a demand that is based on their *a priori* expected utility. The resulting *ex ante* demand function is described by the solid line in Figure 1, D^{ante}. Accordingly, the consumer consumes quantity Q_1.

Now let us assume that the book receives a prize. If the consumer interprets it as a quality signal, she revises her expected utility upwards, and the demand curve also shifts upwards from D^{ante} to D^{post}. The consumer therefore unambiguously increases her consumption from Q_1 to Q_2. However, the consequence of the shift for the consumer surplus depends on the alignment of the prize with her true taste. Specifically, her surplus depends on whether the prize is awarded to a book that she will enjoy more than she initially expected or to a book on which she had correct priors.

If the prize is aligned with the consumer's taste, then the dotted curve is the consumer true demand function. Without the prize, her surplus would have been the sum of Regions A and B. The consumer would have consumed Q_1 but obtained more utility from it than expected. Thanks to the prize, the consumer increased her consumption to Q_2. Her surplus is now given by the whole triangle under the dotted curve, which is the sum of regions A, B, and C. The prize has therefore increased her utility by the dashed triangle C, which is the value of the prize for the consumer.

Figure 1: *The Welfare Effect of a Prize*



Notes: The solid line indicates consumer *ex ante* demand (absent the prize), and the dashed line indicates consumer *ex post* demand (in the presence of the prize). If consumers and experts have similar tastes, consumer surplus is given by $A + B + C$; if their tastes differ, the surplus is equal to $A - D$.

The prize may, however, be poorly aligned with the consumer's taste. In the worst-case scenario, the consumer would have correctly anticipated the utility she will get from the book and her true demand curve indeed corresponds to D^{ante}. If she nonetheless interpreted the prize as signaling a greater utility, she still shifted her demand upwards to D^{post} and increased her demand from Q_1 to Q_2, but this shift was driven by overoptimistic expectations. As a result, the consumer surplus is equal to Region A minus Region D. The prize therefore reduced consumer utility by the dotted triangle D.

In summary, a prize increases the surplus of consumers whose tastes are aligned with the prize and decreases the surplus of consumers whose tastes are not aligned with it. Overall, the welfare effect of the prize is the sum of the variations in the surpluses of all consumers. In a nutshell, it is the sum of all Cs and Ds. It therefore depends on the share of consumers whose tastes are aligned or misaligned with the prize. It also depends on the magnitude of the *ex ante* underestimation of utility by consumers whose tastes are aligned with the prize – the size of their Cs – and on the *ex post* misalignment of the expectations of consumers whose tastes are misaligned with the prize – the size of their Ds.[2]

The impact of the prize on welfare is therefore *a priori* ambiguous. In the following sections, we leverage the specificities of the Booker to estimate its welfare effect.

Figure 1 considers each book individually. In reality, consumers face many books at once. The welfare effect of the Booker might therefore depend on both a substitution effect, whereby consumers switch from non-awarded books to awarded and supposedly better books, and a market expansion effect, whereby consumers increase their total book consumption. In Section 8, we calibrate a model of consumer demand for books where we allow for substitution between books. This makes it possible to study the respective roles of substitution and market expansion in the overall welfare change.

3 A Brief Overview of the Booker Prize

Created in the United Kingdom in 1969, the Booker is one of the most prestigious English-language literary awards (Moseley *et al.* [2019]). The prize is bestowed annually by a jury of five experts to the "best sustained work of fiction written in English and published in the UK and Ireland."[3] The jury members – usually prominent figures on the literary scene (authors, academics, critics, etc.) – change each year and are elected by an advisory committee appointed by the Booker Prize Foundation (Butler *et al.* [2016])

The award is bestowed after several selection stages. From January to July, the judges meet once a month to establish a longlist of 12 to 13 books worthy of winning the prize; in September, the jury announces a shortlist of six books; in October, the winning book is announced. The laureate receives £50,000, while shortlisted authors are awarded £2,500.

Although bestowed by literary experts, the Booker officially aims at awarding the prize to books that will appeal to the widest possible audience. In a 2022 interview, Gaby Wood, Director of the Booker Prize Foundation, stated about the Booker jury: "Essentially what you're looking for is people that are going to read on behalf of the

[2] It is worth noting that Figure 1 focuses only on the intrinsic utility of reading a book. However, readers may also receive extrinsic utility from discussing the book with other readers, which is the basic premise of Adler's [1985] theory of superstars, whereby consumers have an incentive to coordinate on consuming the same cultural products to maximize the probability of being able to discuss them. We do not take that extrinsic utility into account in this paper, and our results must therefore be understood as pertaining to the effect of prizes on the intrinsic utility of purchasing a book.

[3] See https://thebookerprizes.com/the-booker-prize.

general public, but not second guess them." Neil MacGregor, chairman of the 2022 jury, further stated, "We've been looking for books we'd like to recommend to friends."[4]

The key argument of our paper is precisely that the tastes of the jury, or those of its friends, lay at the core of the effect of the prize on welfare. If its tastes are representative of those of the public, the prize will redirect readers to books that they will enjoy. However, if the jury tastes are specific in some way and not aligned with those of readers, the prize may prompt readers to read books they will not enjoy or will enjoy less than the books they would have otherwise read, thereby reducing welfare.

4 Data

To assess the welfare effect of the Booker, we need information on sales, prices, and consumer satisfaction for a representative sample of books. To that end, we constructed a dataset of titles released over the 2015-2021 period by leveraging several sources.[5] The first consists in the titles that were nominated for the Booker during that period (91 titles). Then, we added the titles appearing in the *USA Today best-selling books* ranking (5,865 titles). To avoid having only popular books in our dataset, we supplemented it by including all titles featured in the *Goodreads' yearly book release* lists[6] (6,755 titles) and the titles reviewed by the magazine *Publishers Weekly* (45,303 titles), which consist of both popular and less popular titles. Finally, we collected all the editions of the titles in our dataset, as our data are available at the level of book edition. We ended up with a dataset of 170,941 editions across 58,014 titles.

To collect data on sales and prices, we extracted information on quantities and pre-purchasing characteristics from Amazon.com. Specifically, we observe the daily sales rank, price, average rating, and number of consumer reviews on Amazon of each edition in our dataset from its release date till May 5, 2023. Amazon sales rank is a metric that gauges the sales performance of a product relative to other products listed in the same category. As a result, it moves inversely with actual sales, meaning that a higher rank indicates lower sales. We obtained the Amazon sales ranks and pre-purchasing information of 162,326 editions out of the 170,941 of our initial dataset.

[4] See https://thebookerprizes.com/the-booker-library/features/what-its-really-like-to-be-a-booker-prize-judge.

[5] We focus on the 2015-2021 period as sales records and pre-purchasing information are not available prior to 2015.

[6] Goodreads is a platform dedicated to book lovers. See https://www.goodreads.com.

Using data from Amazon comes with the main advantage of frequency: By having daily observations, we can both identify the impact of the Booker through a sales-based discontinuity strategy and exploit fine-grained variations in prices across editions and over time to assess their impact on consumer demand. In addition, as Amazon represents 42% of the physical book market and 89% of the e-book market in the US (see footnote 1), our data capture a consequential part of the US book market. On the other hand, the main difficulty of using that data is that we observe sales ranks instead of actual sales, as Amazon does not disclose the latter. This raises two issues. First, it makes our results quantitatively difficult to interpret. Second, it makes it impossible to directly compute the price elasticity of demand of a book and the percentage change in sales induced by the Booker that are needed to calibrate our structural model. However, we can circumvent that difficulty by following Brynjolfsson *et al.* [2003] and Chevalier and Goolsbee [2003]. Specifically, the idea is to estimate a regression that relates the actual sales of an edition to its sales rank on Amazon assuming that this relationship follows a power law. We can implement this method for a small subset of books for which we have true sales.[7]

Finally, to assess whether the Booker prompts consumers to read books they enjoy, we need a measure of consumer *ex post* satisfaction from reading a book. We did so by performing a sentiment analysis on the textual content of the reviews posted on Amazon. A sentiment analysis is a natural language processing technique for extracting the sentiment valence of an opinionated text (Pang and Lee [2008]). It classifies each review as either negative or positive, thereby providing a measure of the reviewer satisfaction from reading a particular book.[8] As an alternative measure of satisfaction, we also used the review star rating (number of stars), whereby a higher rating indicates a higher consumer *ex post* satisfaction. Although ratings do not consider all the subtleties of a textual content, they have the advantage of being a more straightforward

[7] We describe the method in detail in Section 5.2. To obtain data on true sales, we leverage the bestseller lists published by Publishers Weekly (see https://www.publishersweekly.com/pw/nielsen/index.html). Specifically, we collected all the weekly bestseller lists from 2015 to 2023 and we matched them with our data on Amazon ranks. We were able to match 7,379 editions.

[8] To perform our sentiment analysis, we use the "Flair" natural language processing framework (Akbik *et al.* [2019]). Flair offers two main advantages. First, it has been shown to produce very accurate predictions (Lien *et al.* [2022], Villanes and Healey [2023]). Second, the model has been trained on a corpus of movie and product reviews, which means that it is particularly suited to our goal of predicting the sentiment of book reviews on Amazon. In Table B.9 of Appendix B.8, we replicate our results with two other popular sentiment analyzers – TextBlob and VADER (Mahrukh *et al.* [2023]) – and obtain very similar results. Our findings are therefore not driven by the type of model used.

measure as they are not algorithm-based. Our dataset contains all consumer reviews written for the books included in our dataset – that is, 6,683,844 reviews.

Table 1 presents summary statistics for the main variables of interest in our sample, separately for awarded and non-awarded titles. Panel A focuses on the daily pre-purchasing information extracted from Amazon, which are available at the level of editions. It shows that awarded editions are, on average, more expensive, sell more (lower sales rank), and have a higher number of ratings than non-awarded editions. Somewhat more surprisingly, we observe that awarded editions are less well rated. In Panel B, we focus on the individual reviews posted by consumers on Amazon that are available at the title level. We again observe that consumers are more likely to post a negative review for awarded books, both in terms of sentiment and rating. Those findings can be interpreted as suggestive preliminary evidence of lower consumer satisfaction with awarded books.

Table 1: *Summary Statistics*

	(1) Awarded	(2) Non-awarded	(3) Difference
A. Quantities and pre-purchasing information			
Sales rank	505,231	894,182	-388,951***
	(6,099)	(144.063)	(6,101)
Price	17.828	16.626	1.203***
	(0.0403)	(0.00512)	(0.0406)
Number of ratings	7,301	1,091	6,210***
	(59.172)	(1.388)	(59.187)
Star rating	4.0642	4.439	-0.374***
	(0.00212)	(0.0000401)	(0.00212)
Number of daily observations	30,067	81,738,566	
B. Consumer individual reviews			
Negative	0.315	0.175	0.141***
	(0.0576)	(0.00150)	(0.0539)
Positive	0.685	0.825	-0.141***
	(0.0576)	(0.00150)	(0.0539)
Star rating	3.971	4.409	-0.438**
	(0.232)	(0.00460)	(0.217)
Number of observations	8,482	6,673,274	

Notes: The variables and the data sources are described in Section 4. Robust standard errors in parentheses, clustered at the book title level in Panel B. ***Significant at 1% level; **Significant at 5% level; *Significant at 10% level.

5 The Impact of the Booker Prize on Sales

As the Booker will only affect consumer surplus if it attracts readers to consumption in the first place, the first purpose of this section is to determine whether the prize improves a book's sales rank. We then infer the estimates of the price elasticity of a book and the percentage change in sales induced by the Booker that we will use to calibrate the structural model used in Section 8 to gauge the welfare effect of the Booker.

5.1 Identification and Results

To address endogeneity, we elaborate on the method used by Reimers and Waldfogel [2021], who study the effect of reviews, and we implement a discontinuity-based approach. Specifically, the idea is to track the sales of books over time and test whether the attribution of the Booker to a book leads to a jump in its daily sales, conditional on controlling for each edition's unobserved quality through the inclusion of fixed effects. This boils down to estimating the following equation:

$$\begin{aligned} \ln(Rank_{it}) = \lambda \ln(Rank_{i,t-1}) + \kappa \ln(p_{it}) + \boldsymbol{\beta}'\mathbf{X}_{it} + \tau 1(Booker = 1)_{it} \\ + \mu_i + f(P_{it}, B_{it}) + \epsilon_{it}, \end{aligned} \tag{1}$$

where $Rank_{it}$ and p_{it} are the Amazon sales rank and price for edition i on day t, and $\mathbf{X}_{it}$ is a vector of control variables (consumer average rating and number of reviews on Amazon). The term $1(Booker = 1)_{it}$ is an indicator that measures the impact of the Booker from its attribution until the next edition of the prize, that is, a period of around one year. To allow for more flexibility and track the effect of the prize over time, we replace in some specifications the $1(Booker = 1)_{it}$ indicator with six indicators coding six two-months periods: one for 0 to 2 months after the attribution of the prize, another for 2 to 4 months, and so on. Finally, μ_i are edition fixed effects, and $f(P_{it}, B_{it})$ is a flexible functional form (up to a second-order polynomial) that models the number of days P_{it} that has elapsed since the publication of edition i and the number of days B_{it} that has elapsed since the attribution of the Booker.

The results of Equation 1 are presented in Table 2. In Column (1), we model the impact of the Booker using a single indicator. The coefficient of the indicator is significant at the one-percent level and equal to -0.184. Accordingly, winning the Booker decreases a book's sales rank by $100 \times (\exp\{-0.184\} - 1) = 16.8\%$, which means that the prize boosts the number of copies sold.[9]

In Column (2), we perform the same exercise, but this time we focus only on the books that have been nominated for the Booker. The coefficient of the Booker decreases in absolute value but remains large ($\tau = -0.0478$) and significant, showing that the Booker increases the sales of awarded books even compared to the sub-sample of nominated books.

In Column (3), we replace the Booker's unique indicator of Column (1) with the six indicators described above, thereby allowing for more flexibility in the timing of the effect. The coefficients of all indicators are negative, statistically significant at the one-percent level, and decrease over time in absolute value. This indicates that the impact of the Booker declines over time during the year following its attribution. Specifically, the prize reduces a book's sales rank by 36.5% in the first two months following its attribution, whereas the effect drops to 0.0546% eight months later.

Identifying the effect of an edition's price on demand is less straightforward as the variable is continuous, meaning that we cannot apply the discontinuity-based approach used to assess the effect of the Booker. In Column (4), which is our most complete specification, we address this issue by instrumenting the edition's price with the number of sellers offering that edition on Amazon. In line with Reimers [2019], we assume that the number of sellers is a proxy for the ease – and, hence, the cost – of distributing an edition. The exclusion restriction underlying this approach rests on the assumption that the number of sellers is uncorrelated with demand shocks. We believe this assumption to be plausible for two reasons. First, our specification includes edition fixed effects – meaning that we restrict the analysis to variations within each edition – while demand shocks are likely occurring at the title rather than edition level. Second, we observe a negative correlation ($\rho = -0.0279$) between an edition's number of sellers and its price, which means that we can also rule out the concern that higher prices increase the number of sellers. The first stage is reported in Table A.1 of Appendix A and shows that the instrument strongly correlates with price.

The results of the second stage – reported in Column (4) – show that the coefficient of log price is equal to 0.00920 and significant at conventional levels, meaning that a one-percent increase in an edition's price leads to a 0.01% increase in its sales rank. If we turn to the six indicators of the Booker, we see that the effect of the prize is qualitatively and quantitatively similar to the previous estimates.

[9] The coefficient must be transformed to be interpreted, since we estimate a semi-log linear regression.

Table 2: *The Impact of the Booker Prize on Amazon Sales Ranks*

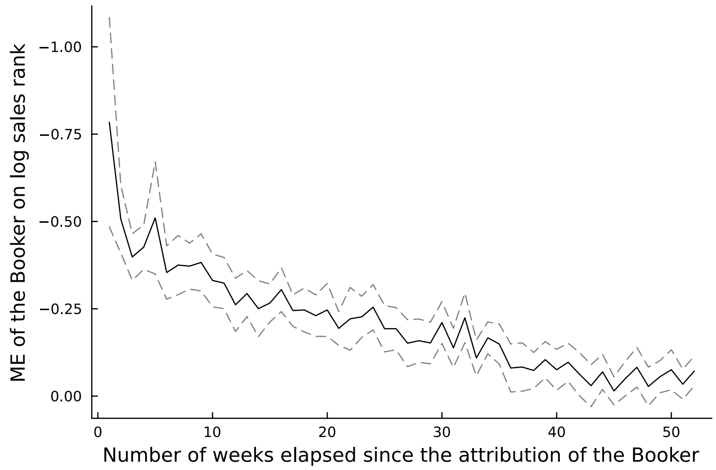
	Outcome: log sales rank			
	(1)	(2)	(3)	(4)
1(Booker=1)	-0.184***	-0.0478***		
	(0.00590)	(0.00563)		
Log price	6.458e-06***	0.00208***	6.460e-06***	0.00920***
	(1.018e-06)	(2.303e-04)	(11.019e-06)	(1.082e-04)
Flexible effect of the Booker				
0-2 month			-0.454***	-0.489***
			(0.0222)	(0.0221)
2-4 months			-0.288***	-0.335***
			(0.012)	(0.011)
4-6 months			-0.220***	-0.229***
			(0.012)	(0.012)
6-8 months			-0.157***	-0.164***
			(0.0106)	(0.0106)
8-10 months			-0.0715***	-0.0744***
			(0.0102)	(0.0102)
10-12 months			-0.0562***	-0.0598***
			(0.00938)	(0.00936)
F Statistics				24,059.7
Sample	All Books	Nominated	All Books	All Books
Adjusted R-squared	0.952	0.976	0.952	0.895
Observations	81,768,633	303,883	81,768,633	80,791,906

Notes: The unit of observation is a day. The model specification follows Equation 1. The dependent variable is an edition's daily Amazon sales rank (in log); a lower sales rank indicates more quantities sold. The term $1(Booker = 1)$ is an indicator that takes value one if the edition is awarded the Booker. The *Flexible effect of the Booker* rows indicate the effect of the Booker for the corresponding number of months following the attribution of the prize. In Column (4), the edition's log price is instrumented by the log number of sellers that offer that edition on Amazon. In each specification, we control for the edition's daily Amazon log sales rank one-day lag, log average rating, and log number of reviews. Each specification also includes edition fixed effects, a flexible control for the number of days that elapsed since the publication of the edition, and a flexible control for the number of days that elapsed since the attribution of the Booker. Robust standard errors are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

To provide a better glimpse of the evolution of the effect of the Booker over time, we ran a similar specification as Equation 1, this time with 52 dummies corresponding to the 52 weeks of the year following the award. Figure 2 summarizes the results of that regression by plotting the coefficients of those dummy variables measuring the marginal effect of the prize on sales ranks over time. Unsurprisingly, the impact of the Booker is strong and sizeable in the first weeks following its attribution, and its effect then decreases over time before becoming indistinguishable from zero after a year.

The results of this section sketch a consistent picture: Being awarded the Booker fosters consumer demand for a book, while an increase in prices curbs it.

Figure 2: *Marginal Effect (ME) of the Booker Prize on Sales over Time*



Notes: Figure 2 is constructed by regressing the log sales rank on 52 week dummies (one for each week following the attribution of the Booker), while controlling for the log sales rank one-day lag, the log average rating, the log number of reviews, edition fixed effects, a flexible control for the number of days elapsed since publication, and the number of days elapsed since the attribution of the Booker. The dashed line reports 95% confidence intervals based on robust standard errors. The y-axis is inversed to reflect the fact that sales ranks move inversely with actual sales.

5.2 Translating Rank Estimates into Quantity Estimates

To compute demand elasticities and quantify the impact of the Booker on sales and welfare, we need to translate the sales rank estimates into sales quantity estimates. We do so by following Brynjolfsson *et al.* [2003] and Chevalier and Goolsbee [2003]. The idea here is to estimate a regression that relates the actual sales of an edition to its sales rank on Amazon by assuming that this relationship follows a power law, that is,

$$Quantity_{iwy} = \sum_{t \in w,y} \mathrm{B} Rank_{ity}^{-\Gamma} + \epsilon_{iwy}. \tag{2}$$

Here, $Quantity_{iwy}$ is the actual number of copies sold by edition i during week w in year y, $Rank_{ity}$ is the sales rank of edition i on day t in year, and η_{iwy} is the error term. Using nonlinear least squares to estimate Equation 2, we find that $\mathrm{B} = 10,320.734$ (498.288) and $\Gamma = 0.346$ (0.0100), where standards errors (in parentheses) are obtained from 100 non-parametric bootstrap draws.

The estimate of Γ can then be used to translate the rank elasticities obtained from Equation 1 into quantity elasticities. The price elasticity of demand for a book is thus given by

$$\epsilon_p = \frac{\partial \ln(Quantity_i)}{\partial \ln(p_i)} = \Gamma \frac{\partial \ln(Rank_i)}{\partial \ln(p_i)} = \frac{\Gamma\kappa}{1-\lambda}, \tag{3}$$

where λ is the coefficient of lagged sales rank and κ the log price's coefficient in Equation 1. Both are estimated in Column (4) of Table 2.[10] The effect of the Booker on sales can similarly be summarized by

$$(\Delta Quantity_i \mid Booker_{ik} = 1) = \frac{\Gamma\tau_k}{1-\lambda}, \tag{4}$$

where k refers to the Booker indicator (0-2 months, 2-4 months, etc.) and τ_k to the associated coefficient estimated in Column (4) of Table 2.

Table 3: *Quantity Effects*

	(1) Baseline	(2) -2SD	(3) +2SD
Price elasticity	-0.0186***	-0.0180***	-0.0192***
	(0.000568)	(0.000544)	(0.000597)
Effect of the Booker			
0-2 month	0.986***	1.065***	0.904***
	(0.0538)	(0.0543)	(0.0533)
2-4 months	0.675***	0.714***	0.636***
	(0.0274)	(0.0290)	(0.0274)
4-6 months	0.462***	0.506***	0.417***
	(0.0286)	(0.0296)	(0.0282)
6-8 months	0.330***	0.369***	0.289***
	(0.0252)	(0.0262)	(0.0258)
8-10 months	0.150***	0.189***	0.110***
	(0.0207)	(0.0212)	(0.0218)
10-12 months	0.121***	0.157***	0.0835***
	(0.0194)	(0.0205)	(0.0203)
Average % effect of the Booker on annual sales	49.850***	54.903***	44.798***
	(2.778)	(3.192)	(2.530)

Notes: The *Price elasticity* row indicates the percentage change in sales with respect to the percentage change in price. The *Effect of the Booker* rows show the percentage impact of the Booker on sales for the corresponding number of months following the attribution of the prize. The last row simulates the average percentage impact of the Booker on annual quantities sold. Figures are based on the coefficients estimated in Column (4) of Table 2. Standards errors are obtained from 100 non-parametric bootstrap draws. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

The results associated with the baseline estimates of Table 2 are reported in Column

[10] As Equation 1 is a partial adjustment model, the derivative of the log rank with respect to the log price is obtained by setting $\ln(Rank_{it}) = \ln(Rank_{i,t-1})$. We therefore have that $\frac{\partial \ln(Rank_i)}{\partial \ln(p_i)} = \frac{\kappa}{1-\lambda}$.

(1) of the upper panel of Table 3. The first noteworthy finding is that the price elasticity of demand is equal to -0.0186, in line with previous estimates (Reimers and Waldfogel [2017, 2021]) which also report an inelastic demand for books. The second set of findings concerns the effect of the Booker on sales, which is sizeable. For example, in the first two months following its attribution, the Booker increases sales by $100 \times (\exp\{0.986\} - 1) = 168\%$.

The parameters B and Γ also allow us to convert each edition's daily rank into daily quantities sold to simulate the effect of the Booker on sales in each calendar year. From the power law relationship between sales and ranks described above (see Equation 2), it follows that the sales of edition i on day t in year y is equal to

$$q_{ity} = \frac{\mathrm{B}}{\exp\left\{\Gamma \ln(Rank_{ity})\right\}}. \tag{5}$$

We can also define the counterfactual sales of edition i – i.e., its sales absent the Booker – by subtracting from its sales the effect of the Booker as defined in Equation 4. That is,

$$q^c_{ity} = \frac{\mathrm{B}}{\exp\left\{\Gamma \ln(Rank_{ity}) - \Gamma \sum_{k=1}^{6} \frac{\tau_k}{1-\lambda} 1(Booker_k = 1)_{ity}\right\}}. \tag{6}$$

The percentage effect of the Booker on annual sales is then obtained by summing q_{ity} and q^c_{ity} over all days of the year and by comparing actual annual sales with counterfactual annual sales:

$$\%\text{ effect of the Booker on the sales of year } y = \frac{\sum_{t \in y} q_{it}}{\sum_{t \in y} q^c_{it}} - 1. \tag{7}$$

We can average Equation 7 over years to obtain an average annual effect.

Column (1) of the lower panel of Table 3 reports the results of this exercise. It shows that the Booker raises annual sales by 50%, on average, which corresponds to an increase of around 216,000 copies per year.

Our computations of quantity effects are based on the price and Booker indicators coefficients estimated in Column (4) of Table 2, and on the parameters B and Γ estimated above. To assess the sensitivity of our results, we compute quantity effects based on ±2 standard deviations of our initial estimates of the coefficients of price, Booker, B, and Γ. The results, reported in Columns (2) and (3) of Table 3, reassuringly have very low sensitivity to the coefficient estimates.

6 The Impact of the Booker Prize on Consumer *Ex Post* Satisfaction

Section 2 shows that the Booker could be welfare increasing or decreasing depending on the distance between the tastes of consumers and those of the jury. To assess the direction of the change in consumer surplus, we investigate how the Booker affects consumer *ex post* satisfaction from reading a book, which we measure with both the sentiment valence and the star rating of the reviews posted on Amazon.

Because we observe the date on which each review was posted, we can compare how sentiments and ratings for awarded and non-awarded books change after the attribution of the prize. In practice, our approach can be summarized by the following difference-in-differences specification:

$$y_{ijt} = \tau 1(Booker = 1)_j \times Post_t + \beta Post_t + \lambda_j + f(P_{it}, B_{it}) + \epsilon_{ijt}, \tag{8}$$

where the variable y_{ijt} is the sentiment valence or star rating of review i for book j in period t (before or after the attribution of the Booker). The sentiment can take two values – zero when negative and one when positive – and the rating discretely ranges from one to five stars. The terms $1(Booker = 1)_j$, $Post_t$, and λ_j are, respectively, an indicator equal to one if book j is awarded the Booker, an indicator set to one for all reviews posted after the attribution of the Booker, and book fixed effects.[11] Finally, the flexible function $f(P_{it}, B_{it})$ models the number of days P_{it} that elapsed between the publication of review i and the publication of the book, and the number of days B_{it} that elapsed between the publication of review i and the attribution of the Booker.

The main parameter of interest is τ, which measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. Therefore, under the assumption that awarded and non-awarded books would have followed the same trend in the absence of the Booker, τ measures the causal impact of the Booker on reviews' sentiments and ratings.

In our baseline specifications, we cluster the standard errors at the book title level to allow for arbitrary dependence between reviews of the same title. To match the effect of the Booker defined in Equation 1 and depicted in Figure 2, we restrict the sample

[11] In specifications where the inclusion of book fixed effects is impossible due to perfect multicollinearity with the treatment, we instead use fixed effects for the book nomination status (0 = not nominated; 1 = longlisted; and 2 = shortlisted).

to reviews posted in the period ranging from the awarding of the previous prize edition to the awarding of the following edition.

6.1 Difference-in-Differences Estimates

The difference-in-differences estimates are presented in Table 4. In Columns (1) and (2), we perform a preliminary validation test by looking at the effect of the Booker on the sentiment valence and rating of the reviews written before the attribution of the prize. For both outcomes, the Booker dummy is statistically insignificant. Accordingly, there are no pre-existing differences between awarded and non-awarded books in terms of consumer satisfaction, which provides a first set of evidence regarding the validity of the parallel trend assumption underlying our approach. In Section 6.2, we provide additional evidence on the parallel trend and show that we obtain similar results even when using methods that do not rely on this assumption.

Table 4: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction*

	Before the Booker		Diff-in-Diff	
	(1)	(2)	(3)	(4)
Outcome	Sentiment	Rating	Sentiment	Rating
1(Booker=1)	0.0800	0.258		
	(0.0777)	(0.295)		
1(Booker=1)×Post			-0.0313**	-0.151***
			(0.0137)	(0.0439)
Outcome mean	0.833	4.431	0.825	4.409
Observations	3,476,096	3,476,096	6,681,756	6,681,756

Notes: The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). Columns (1) and (2) report the results of a validation test where we look at the effect of the Booker on the sentiment and rating of the reviews written before the attribution of the prize. The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker. Columns (3) and (4) report the difference-in-differences estimates. The model specification follows Equation 8. The term *Post* is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. In Columns (1) and (2), book fixed effects are replaced with fixed effects for the book nomination status to avoid perfect multicollinearity with the treatment. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

The last two columns of Table 4 focus on the difference-in-difference estimates.

In Column (3), we use the review sentiment valence as the first measure of consumer satisfaction. The coefficient of $1(Booker = 1) \times Post$ is equal to -0.0313 and is significant at the five-percent level. Accordingly, when a book receives the Booker, the probability that a consumer posts a negative review increases by 3.1 percentage points. In Column (4), we use the review star rating as our second measure of consumer satisfaction. In this case too, $1(Booker = 1) \times Post$ bears a negative coefficient significant beyond conventional levels, meaning that consumers give, on average, a lower rating to awarded books. Specifically, as the coefficient of the diff-in-diff variable is equal to -0.151 and the average rating of an awarded book before the attribution of the prize is 4.060, the Booker leads to an average star decrease of about 3.7%.

The results of Table 4 show that the Booker leads to a deterioration of the assessment of books by consumers. This suggests that the prize redirects consumers toward books they do not enjoy, thereby negatively affecting their surplus.

6.2 Robustness Checks

Parallel Trend Assumption

The difference-in-differences estimates rest on the assumption that absent the Booker the sentiments and ratings for awarded and non-awarded books would have followed a parallel trend. This assumption is plausible in our case for at least two reasons. First, we observe no pre-Booker differences between winners and non-winners, both in terms of consumer satisfaction (Columns (1) and (2) of Table 4) and other observable characteristics (Figure B.1 in Appendix B). Second, moving away from diff-in-diff models, we implement two alternative approaches that do not rely on the parallel trend assumption for identification. We start by conducting a regression discontinuity in time, which is an application of the canonical regression discontinuity framework where time is used as the running variable (see Appendix B.3 for a description of the method; Hausman and Rapson [2018]). We then use an instrumental variable strategy where we generate a set of internal instruments by either using the heteroskedasticity of the errors (Lewbel [2012]) or exploiting higher order moments (we provide more details on each approach in Appendix B.4; Lewbel [1997]). The results are quantitatively and qualitatively similar to the baseline, which bolsters our confidence in the robustness of our diff-in-diff estimates.

Inference

In our baseline specifications, we use cluster-robust standards errors to account for serial correlation (Bertrand *et al.* [2004]). However, such an approach may yield confidence intervals that have poor coverage when the treatment is skewed or when there are few treated clusters, as it is in our framework, since the number of treated clusters is small relative to the number of control clusters (Roodman *et al.* [2019]). We address that concern in three ways. First, we construct confidence intervals using the subcluster wild bootstrap (Columns (1) and (2) of Table B.6 in Appendix B.5). With few treated clusters, this method has been shown to improve on the cluster-robust variance estimator (Roodman *et al.* [2019]). Second, following the recommendations of Imbens and Kolesár [2016], we compute bias-corrected standard errors that are adjusted for small and skewed samples by applying Bell and McCaffrey's [2002] degree-of-freedom correction. We then use those bias-corrected standard errors to construct our confidence intervals (Columns (3) and (4) of of Table B.6). Finally, we restrict the sample to consider only nominated books, which considerably reduces the asymmetry between the number of treated and untreated clusters (Table B.7 in Appendix B.6). The results sketch a similar picture as the baseline.

Changes in the Population Composition of Reviewers

Our diff-in-diff estimates may capture a change in the composition of the population of reviewers if the Booker attracts reviewers who are more likely to write a negative review than pre-Booker reviewers. In that case, $1(Booker = 1) \times Post$ would reflect the fact that individuals who buy a Booker are simply more inclined to leave negative feedback regardless of their satisfaction with the book – for example, because they are more critical readers – rather than an effect of the Booker on satisfaction. We address that concern by leveraging the fact that our sample includes individuals who, in addition to having reviewed an awarded book, also wrote reviews for non-awarded books. This allows us to exploit within-reviewer variations and see whether the same reviewer rated awarded books more negatively than non-awarded ones. We find that it is the case, both in terms of sentiment and rating, as shown in Table B.8 of Appendix B.7. This lends support to the interpretation of our baseline findings in terms of lower satisfaction for awarded books.

7 Mechanism: A Matter of Taste

The theoretical section suggests that the negative welfare impact of the prize may be driven by a misalignment between the tastes of the jury and those of consumers. However, the finding could also be driven by alternative, potentially concurrent mechanisms. Specifically, one could argue that the jury and consumers have similar tastes but that the prize disappoints consumers because it raises expectations that the book subsequently does not meet (Rossi [2021]). Alternatively, consumers that receive utility from exclusiveness (Leibenstein [1950]) may dislike a book due to its increased popularity regardless of its intrinsic quality and of their tastes.

In this section, we provide three series of tests to support our initial interpretation that prizes deteriorate reviews due to a mismatch between the prize and the tastes of consumers. In the first, we focus on the Goodreads Choice Award for Fiction, which is a prize bestowed by a jury of laypeople. This prize should go to books that are closer to the tastes of consumers than a prize that is awarded by practitioners, like the Booker. If the mechanism that we emphasize is at work, the Goodreads prize should result in less dissatisfaction. In a second series of tests, we exploit the variations in the composition of jury of the Booker across editions. The idea behind this approach is that variations in the composition of the jury may result in variations in the proximity of its tastes with those of consumers and, hence, in the impact of the prize on satisfaction. Finally, we provide direct survey evidence on the reactions of readers to awarded books.

7.1 The Effect of a Prize Awarded by Laypeople: the Goodreads Choice Award

The Goodreads Choice Award for Fiction is a popular prize bestowed by the users of the website Goodreads. Like the Booker, the Goodreads prize adds visibility to the awarded book and signals quality. Unlike the Booker, the Goodreads prize is awarded by several hundred thousand Goodreads users, who are laypeople whose tastes are plausibly closer to those of the average reader than are those of the Booker jury. If the negative effect of the Booker is due to the misalignment of the tastes of its jury with those of the public, the Goodreads prize should accordingly not affect the sentiment of reviews. We therefore expect the Goodreads prize to foster sales, like the Booker, but not to negatively affect reviews, unlike the Booker.

We identify the impact of the Goodreads Choice Award for Fiction on sales and on consumer satisfaction using the same identification approaches as for the Booker; that is, we rely on Equation 1 to assess the effect on sales and Equation 8 to investigate the effect on satisfaction. The results, presented in Table 5, confirm our hypothesis: The Goodreads prize indeed boosts sales, resulting in a lower sales rank. By contrast, it has no effect on review sentiment and rating. These findings are consistent with a model where prizes attract readers to consumption, regardless of the composition of their jury, but where consumers may lose utility if their tastes are too far from those of the jury. Those findings are in line with those of Lagios and Méon [2023], who also contrasted the effect on sales and reviews of two French literary prizes awarded from the same list – one by a jury of experts and the other by highschoolers.

Table 5: *The Effect of the Goodreads Choice Award on Sales and Satisfaction*

		Consumer Satisfaction	
	(1)	(2)	(3)
Outcome	Sales Rank	Sentiment	Rating
Effect of the prize	-0.131***	-0.00530	0.00332
	(0.00579)	(0.0146)	(0.0330)
Observations	81,762,885	6,847,207	6,847,207

Notes: The dependent variable is reported at the top of each column. *Sales Rank* refers to the Amazon sales rank of the book, *Sentiment* to the sentiment valence of the review (negative or positive), and *Rating* to its star rating (number of stars). In Column (1), the unit of observation is a day. The specification follows Equation 1 and includes controls for the edition's daily Amazon log sales rank one-day lag, log average rating, and log number of reviews. We also include edition fixed effects, a flexible control for the number of days that elapsed since the publication of the edition, and a flexible control for the number of days that elapsed since the attribution of the prize. Robust standard errors are reported in parentheses. In Columns (2) and (3), the unit of observation is a review. The model specification follows Equation 8. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the prize. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

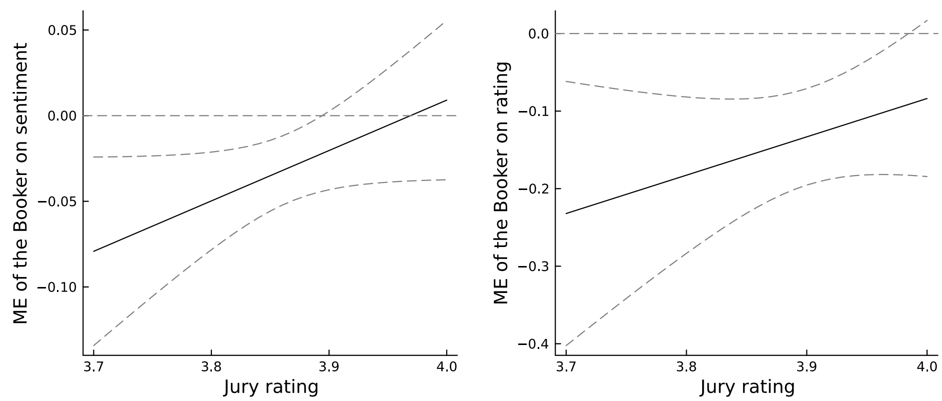
7.2 The Booker Jury's Changing Composition and Consumer Satisfaction

The Booker jury changes each year, which allows us to assess how its composition affects consumer reviews. We focus on two dimensions: the ability of the jury to select a book that consumers will like and the cultural proximity of the jury with the readers.

7.2.1 The Jury's Ability to Select Books Readers Will Enjoy

Experts are more likely to redirect readers towards books they will enjoy if they have the same tastes as the average reader or if they can correctly predict her tastes. As the members of the Booker jury are often writers themselves, we can measure their ability to select a book that readers will enjoy by using the readers' rating of the books the members of the jury have themselves authored. If judges can write books that appeal to consumers, then they might be more likely to give the award to a book that consumers will also enjoy. We measure that capacity for the jury as a whole and refer to it as jury rating. To construct the jury rating, for each awarding of the Booker, we collected all books written by the members of the jury and averaged their readers' ratings. We therefore have one jury rating per Booker edition. We expect editions where the jury's books are less well noted to have a larger negative impact on review sentiment and rating.

Figure 3: *Marginal Effect (ME) of the Booker Prize on Consumer Satisfaction as a Function of the Jury Rating*



Notes: The unit of analysis is a review. The plots are obtained by conditioning the effect of the Booker on the jury rating in Equation 8 (see footnote 12). The left-hand side uses the review sentiment as dependent variable (negative or positive), while the right-hand side uses the review rating (number of stars). Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. The dashed lines indicate 90% confidence intervals based on standard errors clustered at the book title level. The raw coefficients of the model are reported in Appendix C.

We test that hypothesis by conditioning the effect of the Booker on the jury rating in

Equation 8.[12] The results are summarized in Figure 3, which plots the marginal effect of the Booker on review sentiment (left-hand side) and rating (right-hand side) as a function of the jury rating. For both measures of consumer satisfaction, the absolute effect of the Booker diminishes when the jury rating increases, and the effect even becomes statistically insignificant when the average rating of the books written by the jury is high enough. In other words, when the jury's ability to select a suitable book for the average reader is high – because they have the same tastes or because they can accurately predict them – Booker-award-winning books stop disappointing audiences.

7.2.2 The Jury's Cultural Proximity with the Readers

Jury members may differ not only in terms of their capacity to write books that sell but also in terms of cultural proximity with the readers. The greater the proximity, the closer the jury's tastes are likely to be to those of readers. If the negative effect of the Booker on consumer satisfaction is driven by the distance between the tastes of the jury members and those of readers, then it should be smaller when the jury is more representative of the general population.

We measure the representativeness of the jury with three easily observable socio-demographic characteristics: the age, country of birth, and education level of its members. Specifically, for each edition of the Booker, we compute the jury's age dispersion, the share of judges born outside England, and the share of judges with a postgraduate degree. A jury exhibiting a higher age dispersion is likely to cater to the tastes of more age groups. Also, as the outcome of the Booker is covered worldwide, a higher share of members born outside England is likely more representative of consumers. Conversely, we except editions with a higher share of postgraduates to be more disconnected from the average reader's tastes and thus to affect reviews more negatively.

[12] Specifically, we extend Equation 8 by including the variable $\textit{Jury Rating}_e$, which represents the jury rating of edition e, as well as its interactions with $1(Booker = 1)_j \times Post_t$ and $Post_t$, so as to estimate the following regression:

$$\begin{aligned} y_{ijt} = &\beta_1 1(Booker = 1)_j \times Post_t + \beta_2 Post_t + \beta_3 \textit{Jury Rating}_e \\ &+ \beta_4 1(Booker = 1)_j \times Post_t \times \textit{Jury Rating}_e + \beta_5 Post_t \times \textit{Jury Rating}_e \\ &+ \lambda_j + f(P_{it}, B_{it}) + \epsilon_{ijt}. \end{aligned}$$

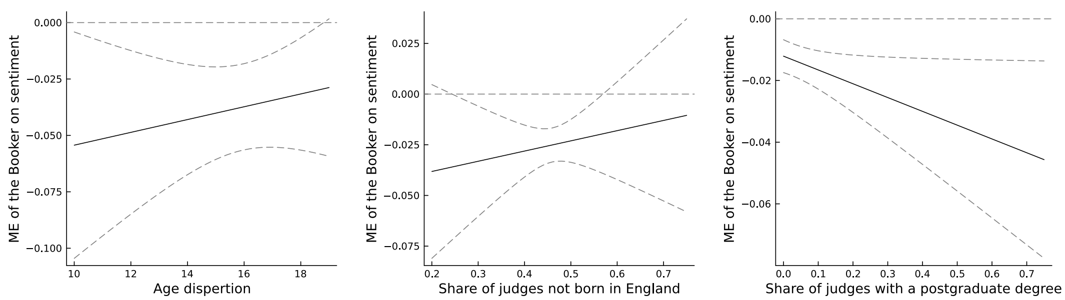
We are interested in estimating the conditional marginal effect of $1(Booker = 1)_j \times Post_t$ on y_{ijt}, that is:

$$\left[\Delta y_{ijt} \mid 1(Booker = 1)_j \times Post_t = 1\right] = \beta_1 + \beta_4 JuryRating_e.$$

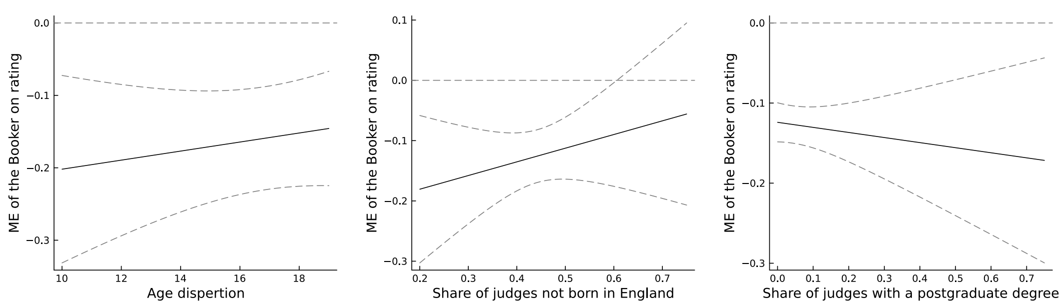
Again, we test the hypothesis by interacting the effect of the Booker with each characteristic of the jury in Equation 8 (see footnote 12). Figure 4 plots the marginal effect of the Booker on review sentiment (Panel A) and rating (Panel B) against our three measures of cultural proximity. Both panels show that the higher the representativeness of the jury, the lower the negative impact of the Booker on reviews sentiment and rating, hence on consumer satisfaction.

Figure 4: *Marginal Effect (ME) of the Booker Prize on Consumer Satisfaction as a Function of the Jury Characteristics*

Panel A. Marginal Effect of the Booker Prize on Sentiment



Panel B. Marginal Effect of the Booker Prize on Rating



Notes: The unit of analysis is a review. The plots are obtained by conditioning the effect of the Booker on our three measures of jury representativeness in Equation 8 (footnote 12). Panel A uses the review sentiment as dependent (negative or positive), while Panel B uses the review rating (number of stars). Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. The dashed lines indicate 90% confidence intervals based on standard errors clustered at the book title level. The raw coefficients of the model are reported in Appendix C.

The results of this section clearly indicate that how a prize affects reviews depends on its jury's representativeness and ability to select books that will appeal to consumers. This suggests that the negative impact of the Booker on consumer satisfaction is driven, at least to some extent, by a divergence between the tastes of the jury and those of consumers.

7.3 Survey Evidence

To get a better view of the reactions of readers to prize-winning books, we leverage an online survey dedicated to consumer reading habits, which we conducted between August 21 and September 5, 2023, on Prolific Academic and which involved 1,000 native English speakers living in the US.[13] The survey features several questions on literary prizes as well as questions on respondents' sociodemographic characteristics. In particular, three questions, whose outcomes are reported in Figure 5, connect to the topics addressed in the previous sections.

The first question allows us to gauge the influence of literary prizes. We asked respondents if they agreed with the statement "When a book has been awarded a literary prize, I am more likely to buy it." Respondents could reply on a scale from "strongly disagree" to "strongly agree" and the distribution of their answers is reported in Figure 5a. Although their answers are split, 47.51% of respondents agree or strongly agree with the statement. Accordingly, almost one half of respondents admit that their decision to buy a book is influenced by literary prizes. In another question, whose results are reported in Appendix D, we asked respondents, "What makes you want to buy a particular book?" 22.66% of them consider "the literary prize(s) it has received" to be either important or very important.[14] Moreover, 58.23% of respondents agree or strongly agree with the statement, "If I hesitate between two books, I am more likely to buy the one which has received a literary prize", which suggests that prizes not only affect the quantity of books sold but also affect which books consumers buy. Overall, those findings confirm the influence of prizes on sales.

We then asked respondents: "How often have you felt disappointed by a book that had been awarded a prize?" Their answers, reported in Figure 5b, show that one half of them report having been disappointed sometimes. This figure increases to 60.6% when including those who have often been disappointed. In addition, 52.49% of respondents agree or strongly agree with the statement that "[they] usually expect the awarded books [they] read to be better than what they actually are." Those findings echo the

[13] Prolific Academic is a crowdsourcing platform dedicated to academic research and other endeavors.

[14] The answers to all the questions that we discuss in this section but are not plotted in Figure 5 are reported in Appendix D.

negative effect of the Booker on the sentiment and rating of online reviews reported in the previous sections.

Figure 5: *Outcome of the Online Survey on Reading Habits*

Figure 5a

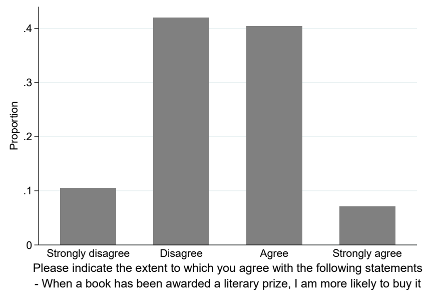


Figure 5b

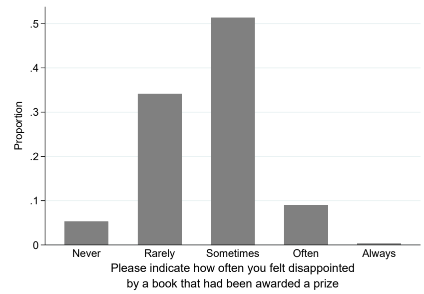
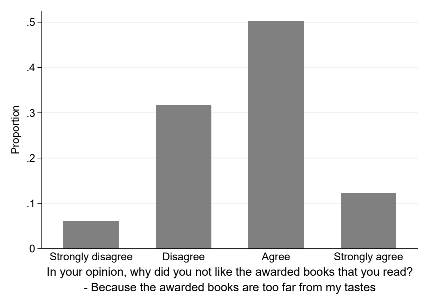


Figure 5c



Notes: The unit of analysis is a survey respondent.

We also asked respondents who reported having been disappointed by awarded books to assess the possible reasons for their disappointment ("In your opinion, why did you not like the awarded books that you read?") and to report their agreement with the statement "because the awarded books are too far from my tastes." 62.37% agree or strongly agree with that statement (Figure 5c), in line with the contention that the tastes of jury members are misaligned with those of most readers.

However, the most striking pieces of evidence appeared when we asked respondents to "please briefly describe [their] experience of reading awarded books". Although anecdotal, their replies confirm a mismatch between their preferences and those of jury members. Many feature the adjective "boring". Many were more explicit, as demonstrated in this admittedly subjective but tasty selection: "Booker Prize winners are usually too 'literary' for me to enjoy." Another respondent wrote, "[T]he awarded books are often corny and robotic." More to the point, some respondents explicitly explain their dissatisfaction by a gap between their tastes and those of the jury: "I tend to find that these books are less accessible, and l feel that the Jury of prizes are disconnected from what I like." Even more pointedly: "I think awards are given by a small group of people who have specific tastes, and chances are my tastes are not similar to the people who gave the award." Finally: "The taste of judges are not my tastes. Its [sic] all subjective and awards are only good for marketing".

Overall, the evidence reported in this sub-section confirms that despite being influenced by prizes when deciding what to read, consumers are often disappointed in prize-winning books. In addition, many of them blame their discontent on a misalignment of the tastes of jury members with theirs, in line with our theoretical contention and the estimated effect of the Booker on reviews.

8 The Effect of the Booker Prize on Consumer Welfare: A Structural Approach

In this section, we quantify the welfare loss induced by the prize (the D triangle in Figure 1) by calibrating a structural model of demand for books. Specifically, our approach consists of simulating consumer surplus in a counterfactual world where the Booker does not exist and comparing it with consumer surplus in the status quo where the Booker does exist.

8.1 Consumer Demand and Surplus

We model consumer demand for books by using a one-level nested logit model (see Berry [1994], Train [2015], Aguiar and Waldfogel [2018], Reimers and Waldfogel [2021]). Such a model allows for substitution between books and for consumers to differ in their reading tastes. Define $\mathcal{J}_t$ as the set of books available at time t and j as the book index. Each consumer makes a discrete choice between purchasing a book from the choice set $\mathcal{J}_t$ or consuming the outside good that consists in not buying a book from the choice set; consumers therefore face $\mathcal{J}_t + 1$ options. Omitting the time subscript for convenience, the utility that consumer i expects to get from choosing book j, which we label "decision utility" following Kahneman [1994], is given by

$$\widetilde{U}_{ij} = \tilde{\delta}_j + \zeta_i + (1 - \sigma)\epsilon_{ij}, \tag{9}$$

where $\tilde{\delta}_j$ is the mean utility consumer i expects to get from purchasing book j and $\sigma \in [0, 1)$ measures the degree of substitution across books. As σ approaches one, books become perfect substitutes for one another, and the entry of an additional book cannibalizes demand for existing books, resulting in a complete business-stealing effect and no market expansion. When $\sigma = 0$, the model collapses to a standard logit in which books are imperfect substitutes and entry leads to an increase in the total number of books read (market expansion). The nested logit model allows for two idiosyncratic taste shock components: ζ_i, which captures consumer i's idiosyncratic tastes for reading books and is common across all books, and ϵ_{ij}, which represents consumer i's idiosyncratic taste toward book j. As shown by Cardell [1997], if ϵ_{ij} is distributed extreme value, then $\zeta_i + (1 - \sigma)\epsilon_{ij}$ is also extreme value distributed.

Our welfare analysis rests on the comparison of consumer surplus under two scenarios: the status quo in which consumers rely on the Booker as pre-purchasing information and a simulated counterfactual in which the Booker does not exist. Specifically, we define the decision mean utility of book j in the status quo, $\tilde{\delta}_j$, and in the counterfactual, $\tilde{\delta}^c_j$, as

$$\tilde{\delta}_j = -\alpha p_j + \omega_j + \xi_j \quad \text{(status quo)} \tag{10}$$

$$\tilde{\delta}^c_j = -\alpha p_j + \xi_j \quad \text{(counterfactual)} \tag{11}$$

where p_j represents the book price, ω_j captures the positive signal of receiving the Booker on expected utility, and ξ_j is a vector of unobserved demand shifters.

In the nested logit demand model, the decision mean utility $\tilde{\delta}_j$ can also be expressed in terms of market shares. Normalizing the mean utility of the outside good to 0, we have

$$\tilde{\delta}_j = \ln(s_j) - \ln(s_0) - \sigma \ln\left(\frac{s_j}{1 - s_0}\right), \tag{12}$$

where $s_j = q_j/M$ and $s_0 = 1 - Q/M$. The term s_j refers to the market share of book j, s_0 to the market share of the outside good, M to the market size, and $Q = \sum_{j \in \mathcal{J}} q_j$ to the sum of all copies sold of the books in the sample $\mathcal{J}$.

Consumers maximize decision utility but given their imperfect knowledge, they may misperceive the utility they will receive from reading an awarded book. Their decision utility, $\widetilde{U}_{ij}$, may therefore not coincide with the utility they actually experience when consuming the book, which we denote as U_{ij} and refer to as "experienced utility" (Kahneman [1994]). As in Allcott [2013], we define U_{ij} to be the same as decision utility, except that now consumers observe the true quality of an awarded book, which causes them dissatisfaction, as shown by the results of Sections 6 and 7:

$$U_{ij} = \delta_j + \zeta_i + (1-\sigma)\epsilon_{ij}, \tag{13}$$

where

$$\delta_j = -\alpha p_j - \gamma\omega_j + \xi_j. \tag{14}$$

In our baseline and most conservative scenario, we set $\gamma = 0$, meaning that consumers' experienced utility is equal to their decision utility absent the Booker. This approach provides a lower bound of the effect of the prize on welfare as it is equivalent to assuming that, absent the Booker, consumers have no misperceptions about quality and are the best judges of the utility they will get from purchasing a given book. However, the Booker may also redirect consumers toward books that they end up disliking even more than what they would have thought in the counterfactual, hence resulting in an experienced utility that is even lower than what consumers initially expected. One can capture this by setting $\gamma > 0$. A convenient way to model that dissatisfaction from reading a Booker Prize-winning book is to assume that it is proportional by a factor γ to the utility ω_j consumers were expecting to obtain. Doing so allows us to explore how welfare changes with γ – that is, how welfare varies as dissatisfaction increases.

Given the nested logit demand system, the change in consumer surplus (CS) from the status quo to the counterfactual scenario is given by:

$$\begin{aligned}\Delta CS = \frac{M}{\alpha}\Bigg\{ &\ln\left(1 + \left[\sum_j \exp\left(\frac{\tilde{\delta}_j}{1-\sigma}\right)\right]^{1-\sigma}\right) \\ &- \ln\left(1 + \left[\sum_j \exp\left(\frac{\tilde{\delta}^c_j}{1-\sigma}\right)\right]^{1-\sigma}\right) \\ &+ \sum_j s_j(\delta_j - \tilde{\delta}_j) - \sum_j s^c_j(\delta_j - \tilde{\delta}^c_j)\Bigg\}.\end{aligned} \tag{15}$$

The term s^c_j refers to the market share of book j absent the Booker, which is defined as $s^c_j = \frac{\exp\{\tilde{\delta}^c_j/(1-\sigma)\}}{D_c^\sigma(1+D_c^{1-\sigma})}$, where $D_c = \sum_{j\in\mathcal{J}} \exp\{\tilde{\delta}^c_j/(1-\sigma)\}$ (Berry [1994]). The first part of Equation 15 represents consumer's expected surplus in the presence of the Booker, which is based on her decision utility (that is, the utility she anticipates). The second part reflects consumer's expected surplus absent the prize. The third part is an adjustment to account for the fact that experienced utility may differ from decision utility. The last part is an adjustment that reflects the fact that, in the counterfactual, consumers make decisions based on $\tilde{\delta}_j^c$ (decision utility absent the Booker) but actually obtain a mean utility equal to the mean utility under the status quo, δ_j. In the baseline, $\left[s_j(\delta_j - \tilde{\delta}_j)\right] < 0$ and, as we assume that consumers have no misperceptions absent the Booker, $\left[s_j^c(\delta_j - \tilde{\delta}_j^c)\right] = 0$. When we allow for imperfect knowledge, then the expression $\left[s_j^c(\delta_j - \tilde{\delta}_j^c)\right]$ becomes positive. We provide more details on the derivation of Equation 15 in Appendix E.

The change in net revenues induced by the Booker is given by the following formula:

$$\Delta \textit{Net Revenues} = M\left\{\sum_j p_j s_j - \sum_j p_j s_i^c\right\} \tag{16}$$

8.2 Estimation Procedure

The calibration of Equation 15 requires estimates for the market size M, the substitution parameter σ, the price utility parameter α, and the Booker utility parameter ω_j. We compute them as follows.

The market size M. In line with previous research (Aguiar and Waldfogel [2019], Reimers [2019], Reimers and Waldfogel [2021]), we assume that M is equal to the population size of the country of interest (in our case, the US) times 12. In other words, we assume that each month every American makes a discrete decision between buying a book or consuming an outside good.

The substitution parameter σ. As in Berry [1994], we obtain σ by estimating the following regression $\ln(s_{jt}) - \ln(s_{0t}) = \sigma \ln\left(\frac{s_{jt}}{1-s_{0t}}\right)$, where the variables are defined as above. Since $\ln\left(\frac{s_{jt}}{1-s_{0t}}\right)$ is by construction endogenous, we instrument it by using the standard BLP instrument, which is the number of available titles (e.g., Nevo [2000], Aguiar and Waldfogel [2018, 2019], Reimers [2019], Berry and Haile [2021]). We obtain σ equal to 0.376, confirming that books are imperfect substitutes for one another. We provide more details on the estimation of σ in Appendix F. We also show in Section 8.3 that our welfare estimates are only slightly sensitive to the value of σ.

The price utility parameter α. The nested logit allows us to obtain a consistent estimate of the utility parameter α. Given our modelling assumptions, the market share of each edition is given by $s_j = \frac{\exp\{\tilde{\delta}_j/(1-\sigma)\}}{D^{\sigma}(1+D^{1-\sigma})}$, where $D = \sum_{j\in\mathcal{J}} \exp\{\tilde{\delta}_j/(1-\sigma)\}$

(see Berry [1994]). It follows that the price elasticity of demand can be computed as

$$\hat{\epsilon}_p = -\alpha_j \frac{1}{1-\sigma}\Big(1 - \sigma\frac{s_j}{1-s_0} - (1-\sigma)s_j\Big)p_j. \tag{17}$$

Given $\hat{\epsilon}_p$ that has been estimated in Table 3, s_j, s_0 and p_j that are observed or can easily be computed in the data, and σ that has been derived above, we can solve for α_j for each edition j, and then average it over all editions to obtain α.

The Booker utility parameter ω_j. We estimate the utility parameter ω_j following Reimers and Waldfogel [2021]. In the empirical approach of Section 5.2, we have identified the impact of the Booker on sales by comparing a book's actual sales q_j with its sales absent the Booker q_j^c (i.e., in the counterfactual). That is, $\ln\left(\frac{q_j}{q_j^c}\right)$. The equivalent in our nested logit model is given by $\ln\left(\frac{s_{j,B}}{s_{j,B}^c}\right) - \ln\left(\frac{s_{j,B'}}{s_{j,B'}^c}\right)$, where $s_{j,B}$ is the sales of awarded books, $s_{j,B}^c$ the sales of awarded books absent the Booker, $s_{j,B'}$ the sales of non-awarded books, and $s_{j,B'}^c$ the sales of non-awarded books absent the Booker.[15] Equating the two expressions, a few lines of algebra show that

$$\omega_j = \ln\left(\frac{q_j}{q_j^c}\right)(1-\sigma), \tag{18}$$

which means that given q_j, q_j^c, and σ, which we know, we can estimate ω_j.

8.3 Results

The results of our welfare analysis are reported in Table 6. We compute the standard errors by using 100 non-parametric bootstrap draws on B, Γ, σ, and the coefficients estimated in Column (4) of Table 2. We first focus on the net revenue generated by the Booker in the book industry to get a glimpse of the impact of the prize on producers, which is the difference between the extra revenues accruing to the awarded book and the loss in the revenues of other books to which readers substitute the awarded one. Our simulation exercise shows that US publishers would be worse off absent the prize as they would have had lower revenues. Specifically, each year, the Booker raises the net book industry revenue by USD688,113 on average.

[15] The expression $\left[\ln\left(\frac{s_{j,B}}{s_{j,B}^c}\right) - \ln\left(\frac{s_{j,B'}}{s_{j,B'}^c}\right)\right]$ can be interpreted as the percentage change in sales for awarded books induced by the Booker with respect to the percentage change in sales of non-awarded books induced by the Booker.

We then turn to the impact of the Booker on consumer surplus. Our welfare computations based on the baseline estimates of α, ω, σ, and γ show that the existence of the Booker decreases consumers surplus by USD135,189 each year. Since the Booker leads to an average increase in book sales of 216,000 copies, this means that each consumer buying a book because it has been awarded experiences a loss in her surplus of USD0.60, which corresponds to 4% of the average price of a book.

Table 6: *The Welfare Effect of the Booker Prize*

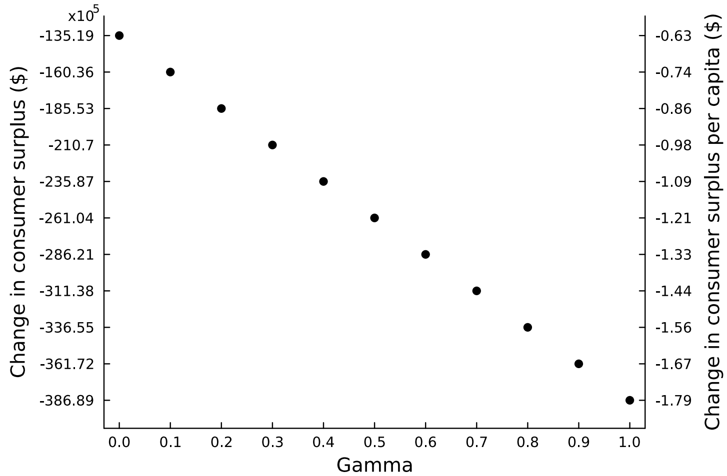
	Effect	SE
Change in net revenues	688.113	52.700
Change in consumer surplus (baseline)	-135.189	8.539
Change in consumer surplus ($\sigma = 0$)	-135.201	9.091
Change in consumer surplus ($\sigma = 0.95$)	-135.171	8.719

Notes: All figures are in thousands of dollars. The change in consumer surplus is computed following Equation 15. The empirical implementation is explained in Section 8.2. Figures are based on the coefficients estimated in Column (4) of Table 2. Standards errors are obtained from 100 non-parametric bootstrap draws.

Table 6 also reports the change in consumer surplus for alternative $\sigma \in [0, 1)$, specifically for $\sigma = 0$ and $\sigma = 0.95$. When $\sigma = 0$, books are imperfect substitutes; when σ approaches one, books become perfect substitutes for one another, and the entry of an additional book cannibalizes demand for existing books. Varying the parameter σ therefore allows us to determine the extent to which the welfare effect of the Booker arises from consumers switching from non-awarded to awarded books or from consumers increasing their total book consumption. As shown by Table 6, our welfare results are insensitive to σ, meaning that our results mainly arise from consumers switching from non-awarded to awarded books, which they expect to enjoy more.

Finally, we assess how our welfare results vary with the parameter γ, which we use in Equation 14 to model consumer dissatisfaction from reading a Booker. The results are documented in Figure 6. The lower bound of the welfare effect of the Booker is our baseline estimates, where we set $\gamma = 0$. When we assume that $\gamma = 1$, to have a symmetric case in which the absolute value of consumer dissatisfaction equals the marginal gain in utility ω_j she was expecting to get when buying the book, we obtain a loss in welfare that is three times larger, and that now accounts for 10.5% of the average price of a book.

Figure 6: *The Welfare Effect of the Booker Prize as a Function of Consumer Dissatisfaction* (γ)



Notes: The change in consumer surplus is computed following Equation 15. The empirical implementation is explained in Section 8.2. Figures are based on the coefficients estimated in Column (4) of Table 2.

9 Conclusion

Expert opinions are ubiquitous and influential, and they are usually believed to help consumers make better-informed decisions. However, they may also draw consumers to products that imperfectly suit consumers. Experts' effect on consumer welfare is therefore *a priori* ambiguous. In line with that argument, we observe that the Booker prize increases sales but decreases the satisfaction of consumers as measured by the sentiment and rating of online reviews. Moreover, we report an array of evidence that the negative effect of the prize on consumer satisfaction is driven by a misalignment between the tastes of the members of the jury of the prize and those of readers. Finally, by calibrating a structural model of demand for books, we estimate a negative and substantial welfare effect of the prize, which questions the role of awards and experts, especially when they concern experience and cultural goods. Those findings imply that the notion of product quality can be misleading when applied to those goods and that the stakes of prizes and experts go beyond signaling the "best" products and may call for a qualification of the way we think about quality.

The argument that we apply to books and prizes equally applies to many types of goods and forms of expert judgments. What matters is that quality be imperfectly observable prior to consumption and that the preferences of experts potentially be misaligned with those of buyers, be they individual consumers, firms, or governments. Our analysis therefore ought to be performed in other industries and forms of expert judgments.

Regardless, one may wonder why readers continue to follow the recommendations of awards despite the suspicion that they may direct them to products that do not correspond to their tastes. One answer may be that awards play the role of coordination devices if consumers get utility from consuming products that are also consumed by others, in line with the mechanism of Adler's [1985] model of superstars and the findings of Lagios and Méon [2023]. Consumers may accordingly trade off intrinsic utility for extrinsic utility, in line with Loeper *et al.* [2014]. We emphasize that the present paper only gauges the intrinsic utility of reading an awarded book. Taking extrinsic utility into account and estimating it would be a natural extension of our analysis and would be necessary to estimate the full effect of awards on social welfare.

That estimation notwithstanding, awards would in any case be superior coordinating devices if they directed consumers to products that give consumers more intrinsic utility. Over time, consumers should favor awards that are closer to their tastes, and misaligned awards should lose influence. Jury members should therefore have an incentive to target the tastes of the median consumer. The persistence of awards that are imperfectly aligned with the preferences of the median consumer is a puzzle and calls for research on the political economy of awards. That research agenda will require a better understanding of the interactions of all the actors of the awards industry: producers, artists, experts, public authorities, and the very organizers of awards themselves. In a nutshell, we need a better understanding of the players, the strategies, and the rules of what French poet Stéphane Mallarmé (1945, cited by Bourdieu [1983]) referred to as "a game".

Appendix

A The Impact of the Booker Prize on Sales – First Stage

Table A.1 presents the first-stage estimates of the 2SLS implemented in Column (4) of Table 2 in the main text.

Table A.1: *First-Stage Estimates – The Impact of the Number of Sellers on Price*

	Outcome: log price	
	Effect	SE
Log number of sellers	-1.147***	0.00740
Adjusted R-squared	0.0520	
Observations	80,791,906	

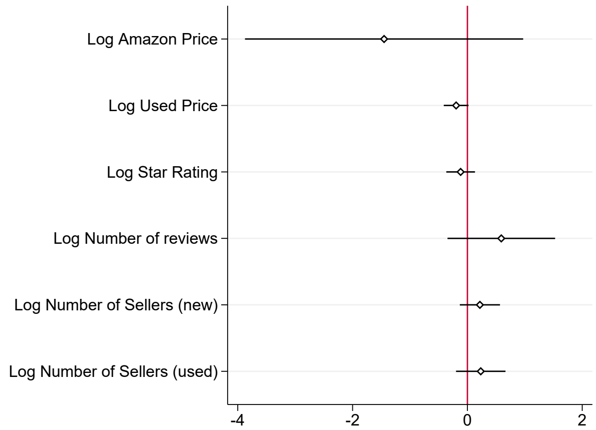
Notes: The unit of observation is a day. The dependent variable is an edition's log price. Log number of sellers refers to the number of sellers that offer that edition on Amazon. The specification includes controls for the edition's daily Amazon log sales rank one-day lag, log average rating, and log number of reviews. It also includes edition fixed effects, a flexible control for the number of days that elapsed since the publication of the edition, and a flexible control for the number of days that elapsed since the attribution of the Booker. Robust standard errors are reported. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

B Robustness Checks for the Effect of the Booker Prize on Consumer *Ex Post* Satisfaction

B.1 Covariate Balance

Figure B.1 provides additional evidence on the parallel trend assumption of the diff-in-diff approach implemented in Section 6. Specifically, we show that there are no systematic differences between winners and non-winners in terms of observable characteristics.

Figure B.1: *Covariate Balance*



Notes: Figure B.1 reports the results of a balance test looking at the effect of the Booker on several editions' observable characteristics before the attribution of the prize. The specification includes a flexible control for the number of days elapsed since publication and the number of days elapsed since the attribution of the Booker, as well as includes fixed effects for the edition nomination status. The variable Number of reviews is in hundreds of reviews. The horizontal black line indicates 90% confidence intervals based on standard errors clustered at the book title level.

B.2 No Control Variables

Table B.1 reports the results of the diff-in-diff approach discussed in the main text but without control variables.

Table B.1: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – No Controls*

	(1)	(2)
Outcome	Sentiment	Rating
1(Booker=1)×Post	-0.0269**	-0.129***
	(0.0120)	(0.0353)
Observations	6,681,756	6,681,756

Notes: The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The model specification follows Equation 8. The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker, and *Post* is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books. Each specification includes book fixed effects. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

B.3 Regression Discontinuity in Time

In this section, we implement a Regression Discontinuity in Time (RDiT) as first alternative to the diff-in-diff approach used in the main text. RDiT is an application of the standard Regression Discontinuity (RD) design framework where time is used as the running variable (Hausman and Rapson [2018]). Following standard practices, we estimate our RDiT regression using a local linear approach where we focus only on observations close to the cutoff (e.g., see Gelman and Imbens [2019]).

The standard RD framework assumes a continuous running variable. However, as time is discrete, the local linear estimator can lead to confidence intervals that have poor coverage (Lee and Card [2008], Kolesár and Rothe [2018]). To address that concern, we draw inference from Armstrong and Kolesár [2018] and Kolesár and Rothe [2018]. Specifically, in addition to conventional CIs, we report "honest" CIs, in the sense that they are valid even in discrete settings. We construct those CIs by using the bounded second derivative (BSD) procedure which requires choosing a constant K that bounds in absolute value the second derivative of the conditional expectation function (Kolesár and Rothe [2018]).

The results of the RDiT are reported in Table B.2 (review sentiment) and Table B.3 (review rating). Each column uses a different K, and for each K a new optimal bandwidth is computed following Kolesár and Rothe [2018]. This allows us to assess the sensitivity of our findings to the choices of both K and the bandwidth. In Column (1), we use a lower bound estimate of K (see the online supplements to Kolesár and Rothe [2018]), while Columns (2) and (3) rely on a K equal to 5 and 15 times the lower bound estimate, respectively. The higher the K, the more conservative the approach. The RDiT estimates confirm that the Booker significantly decreases consumer satisfaction, both in terms of sentiments and ratings.

Table B.2: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – Regression Discontinuity in Time for Review Sentiment*

	Outcome: review sentiment		
	(1)	(2)	(3)
Estimate	-0.0347***	-0.0391***	-0.0457***
	(0.0127)	(0.0120)	(0.0115)
BSD 95% CIs	[-0.0609, -0.00849]	[-0.0738, -0.00444]	[-0.0907, -0.000600]
K	4.005e-07	2.003e-06	6.008e-06
Bandwidth	328	255	184.443
Observations	8,052	7,283	6,145

Notes: Local linear RD estimates with triangular kernel. The unit of observation is a review. The dependent variable is the review sentiment valence (negative or positive). BSD refers to the bounded second derivative procedure which is used to construct "honest" confidence intervals with standard errors clustered at the book level. The approach requires choosing a constant K that bounds in absolute value the second derivative of the conditional expectation function (Armstrong and Kolesár [2018], Kolesár and Rothe [2018]). Column (1) uses a lower bound estimate of K (see the online supplements to Kolesár and Rothe [2018]), while Columns (2) and (3) rely on a K equal to 5 and 15 times the lower bound estimate, respectively. For each K, the optimal bandwidth is computed following Kolesár and Rothe [2018] and minimizes the length of fixed-length two-sided confidence intervals. Each specification includes book fixed effects and a flexible control for the number of days that elapsed between the publication of the review and the publication of the book. Standard errors clustered at the book title level are reported in parentheses.

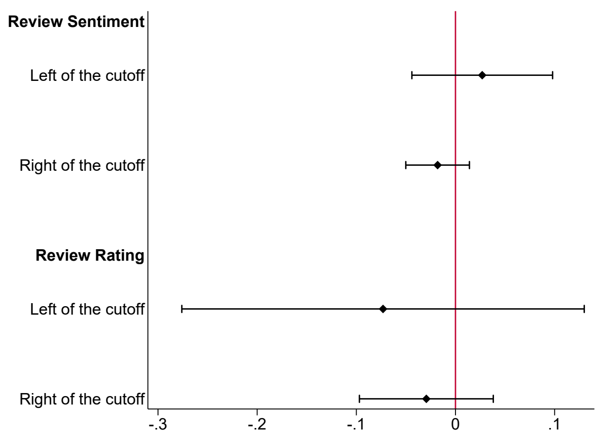
Table B.3: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – Regression Discontinuity in Time for Review Rating*

	Outcome: review rating		
	(1)	(2)	(3)
Estimate	-0.0927**	-0.109***	-0.120***
	(0.0388)	(0.0352)	(0.0431)
BSD 95% CIs	[-0.169, -0.0165]	[-0.249, 0.0314]	[-0.211, -0.0282]
K	2.972e-06	1.486e-05	4.458e-05
Bandwidth	281	211.827	48.265
Observations	7,560	6,637	2,510

Notes: Local linear RD estimates with triangular kernel. The unit of observation is a review. The dependent variable is the review star rating (number of stars). BSD refers to the bounded second derivative procedure which is used to construct "honest" confidence intervals with standard errors clustered at the book level. The approach requires choosing a constant K that bounds in absolute value the second derivative of the conditional expectation function (Armstrong and Kolesár [2018], Kolesár and Rothe [2018]). Column (1) uses a lower bound estimate of K (see the online supplements to Kolesár and Rothe [2018]), while Columns (2) and (3) rely on a K equal to 5 and 15 times the lower bound estimate, respectively. For each K, the optimal bandwidth is computed following Kolesár and Rothe [2018] and minimizes the length of fixed-length two-sided confidence intervals. Each specification includes book fixed effects and a flexible control for the number of days that elapsed between the publication of the review and the publication of the book. Standard errors clustered at the book title level are reported in parentheses.

To show the validity of our RDiT framework, we follow Hausman and Rapson [2018] and perform a placebo test where we investigate the presence of discontinuities at placebo cutoffs – that is, cutoffs where there should normally be no jump. As recommended by Imbens and Lemieux [2008], we implement that test in two steps. First, we divide our sample into two sub-samples, resulting in one sub-sample containing only observations to the left of the cutoff and another sub-sample containing only observations to the right. We then run an RDiT in each of these sub-samples using the median of the running variable as cutoff. The results are reported in Figure B.2 and show no evidence of discontinuities.

Figure B.2: *RDiT: Placebo Cutoffs*



Notes: Non-parametric RD estimates with triangular kernel. The unit of observation is a review. Each specification includes book fixed effects and a flexible control for the number of days that elapsed between the publication of the review and the publication of the book. The horizontal black line indicates 90% confidence intervals based on standard errors clustered at the book title level.

B.4 Instrumental Variable

In this section, we implement a two-stage least squares (2SLS) approach to assess the effect of the Booker on consumer satisfaction. In the absence of a compelling external instrument, we generate internal instruments in two ways: first, using the heteroskedasticity of errors and second, relying on higher moments.

B.4.1 Approach 1. Identification through Heteroskedasticity

Our first approach follows the identification strategy of Lewbel [2012], which leverages the presence of heteroskedasticity in the error term of the first stage to generate a set of internal instruments from the covariates. An important assumption for identification is that the covariance between the instruments and the squared error of the first stage is non-zero. This assumption can be gauged by testing for heteroskedasticity in the first-stage regression. Performing a Breusch-Pagan test, we reject the null hypothesis of constant variance, which bolsters our confidence regarding the validity of our heteroskedasticity-based approach.

The 2SLS estimates are reported in Table B.4. The first noteworthy finding is that the instruments generated by the approach are strong, as shown by the F statistics. The second noteworthy finding is that the results are in line with the diff-in-diff approach discussed in the main text: Both in terms of sentiments and ratings, the Booker decreases consumer satisfaction.

Table B.4: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – 2SLS Estimates with Internal Instruments Based on Heteroskedasticity*

	(1)	(2)
Outcome	Sentiment	Rating
1(Booker=1)	-0.0404**	-0.192***
	(0.0168)	(0.0673)
F Statistics	315.722	315.722
Observations	6,681,756	6,681,756

Notes: 2SLS estimates. The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker. The variable is instrumented by a set of internal instruments generated following the identification strategy of Lewbel [2012], which exploits the presence of heteroskedasticity in the error term of the first stage. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

B.4.2 Approach 2. Identification through Higher Moments

Our second identification strategy relies on the method proposed by Lewbel [1997], where we construct a set of instruments by exploiting higher order moments of the data. The approach rests on the assumption of skewness of the endogenous regressor, which we show to hold by performing the test proposed by D'agostino *et al.* [1990].

The results, reported in Table B.5, are quantitatively and qualitatively similar to the baseline.

Table B.5: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – 2SLS Estimates with Internal Instruments Based on Higher Moments*

	(1)	(2)
Outcome	Sentiment	Rating
1(Booker=1)	-0.0555**	-0.277***
	(0.0263)	(0.0940)
F Statistics	68.768	68.768
Observations	6,681,756	6,681,756

Notes: 2SLS estimates. The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker. The variable is instrumented by a set of internal instruments generated following the identification strategy of Lewbel [1997], which exploits higher order moments of the data. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

B.5 Wild Bootstrap and Bias-Corrected Confidence Intervals

Cluster-robust standards errors may be biased when the number of treated clusters is low or when the treatment is skewed. In this section, we use alternative inference methods that are more robust to such concerns to show the robustness of our baseline results. In Columns (1) and (2) of Table B.6, we use the subcluster wild bootstrap to construct the confidence intervals of our diff-in-diff estimator. With few treated clusters, Roodman *et al.* [2019] show that this method improves on the cluster-robust variance estimator.

In Columns (3) and (4), we compute bias-corrected standard errors that are adjusted for small and skewed samples by applying Bell and McCaffrey's [2002] degree-of-freedom correction (Imbens and Kolesár [2016]). We then use those standards errors to construct our confidence intervals.

In all cases, the results tend to be similar to the baseline.

Table B.6: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – Alternative Inference*

	90% Wild Bootstrap CIs		90% Bias-Corrected CIs	
	(1)	(2)	(3)	(4)
Outcome	Sentiment	Rating	Sentiment	Rating
1(Booker=1)×Post	-0.0313	-0.151	-0.0313	-0.151
	[-.056, -.007]	[-.261, -.042]	[-.071, .008]	[-.281, -.021]
Observations	6,681,756	6,681,756	6,681,756	6,681,756

Notes: The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The model specification follows Equation 8. The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker, and *Post* is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. Columns (1) and (2) report 90% subcluster wild bootstrap confidence intervals based on 999 replications (Roodman *et al.* [2019]). Columns (3) and (4) report 90% bias-corrected confidence intervals by using Bell and McCaffrey's [2002] degree-of-freedom correction (Imbens and Kolesár [2016]). Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker.

B.6 Only Nominated Books

Table B.7 reports the results of regressions that focus only on the reviews of the books that have been nominated for the Booker.

Table B.7: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – Only Nominated Books*

	(1)	(2)
Outcome	Sentiment	Rating
1(Booker=1)×Post	-0.0296*	-0.116***
	(0.0161)	(0.0386)
Observations	42,790	42,790

Notes: The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The model specification follows Equation 8. The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker, and *Post* is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

B.7 Changes in the Population Composition of Reviewers

Our baseline results may be driven by the fact that post-Booker reviewers have characteristics that make them more likely to leave a negative review than pre-Booker reviewers. In Table B.8, we tackle that concern by exploiting within-reviewer variations. Specifically, we focus on reviewers who wrote a review for both awarded and non-awarded books and compare the sentiment and rating of the reviews of awarded and non-awarded books. The results show that awarded books are less well rated than non-awarded books, which supports the interpretation of our baseline findings in terms of lower satisfaction for awarded books.[16]

Table B.8: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – Exploiting Within-Reviewer Variations*

	(1) Awarded	(2) Non-awarded	(3) Difference
Sentiment	0.442	0.665	-0.223***
	(0.0386)	(0.00457)	(0.0364)
Rating	3.239	3.809	-0.570***
	(0.121)	(0.0104)	(0.113)
Observations	774	11,755	

Notes: *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

B.8 Alternative Sentiment Analyzers

In the main text, we use the Flair framework to predict the sentiment of each review. To make sure that our results are not driven by this specific model, we measure the review sentiment with two alternative sentiment analyzers: TextBlob (Loria *et al.* [2018]) and VADER (HHutto and Gilbert [2014]. The conclusions remain unchanged, as shown by Table B.9.

[16] The implementation of this test required us to run a new phase of review scraping to collect the unique ID of each reviewer, as that piece of information was not collected when we initially scraped the data to construct our baseline dataset in the main text. However, between these two scraping phases, Amazon implemented a limit of one hundred to the number of reviews that are shown in the review section. If we filter reviews by star rating, this means that the maximum number of reviews that can be now collected for a book is 500, or 100 per star rating. Above that number, reviews are simply "lost". Because of this limitation, we are able to recover the reviewer unique ID for roughly 50% of the reviews in our dataset.

Table B.9: *The Impact of the Booker Prize on Consumer Ex Post Satisfaction – Only Nominated Books*

	Outcome: review sentiment	
	(1)	(2)
Sentiment Analyzer	TextBlob	VADER
1(Booker=1)×Post	-0.0642***	-0.0407***
	(0.0164)	(0.0124)
Outcome mean	1.810	1.749
Observations	6,681,756	6,681,756

Notes: The unit of observation is a review. The dependent variable is the review sentiment, which refers to the sentiment valence of the review (negative, neutral, or positive). The model specification follows Equation 8. The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker, and *Post* is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

C Mechanisms

In this section, we report the raw coefficients obtained when estimating the interaction effects presented in Sections 7.1 and 7.2 in the main text.

Table C.1: *Interaction between the Booker Prize and the Jury Rating – Raw coefficients*

Outcome	(1) Sentiment	(2) Rating
1(Booker=1)×Post	-1.168	-2.061
	(0.730)	(1.879)
1(Booker=1)×Post×Jury Rating	0.294	0.494
	(0.189)	(0.482)
Observations	6,681,756	6,681,756

Notes: The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The model specification follows Equation 8 where the variable *Jury Rating* and its interactions with $1(Booker = 1) \times Post$ and *Post* are included (see footnote 12 in the main text). The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker, and *Post* is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. The variable *Jury Rating* refers to the average readers' rating of the books the members of the jury have themselves authored. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

Table C.2: *Interaction between the Prize Booker and the Jury's Age Dispersion – Raw coefficients*

Outcome	(1) Sentiment	(2) Rating
1(Booker=1)×Post	-0.0827	-0.264
	(0.0765)	(0.168)
1(Booker=1)×Post×Age Dispersion	0.00284	0.00624
	(0.00474)	(0.00961)
Observations	6,681,756	6,681,756

Notes: The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The model specification follows Equation 8 where the variable *Age Dispersion* and its interactions with $1(Booker = 1) \times Post$ and *Post* are included (see footnote 12 in the main text). The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker, and *Post* is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. The variable *Age Dispersion* refers to the jury's age dispersion. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

Table C.3: *Interaction between the Prize Booker and the Share of Judges Not Born in England – Raw coefficients*

Outcome	(1) Sentiment	(2) Rating
1(Booker=1)×Post	-0.0482	-0.226*
	(0.0455)	(0.129)
1(Booker=1)×Post×Share of Judges Not Born in England	0.0503	0.227
	(0.0983)	(0.286)
Observations	6,681,756	6,681,756

Notes: The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The model specification follows Equation 8 where the variable *Share of Judges Not Born in England* and its interactions with $1(Booker = 1) \times Post$ and $Post$ are included (see footnote 12 in the main text). The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker, and $Post$ is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. The variable *Share of Judges Not Born in England* refers to the share of judges born outside England. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

Table C.4: *Interaction between the Prize Booker and the Share of Judges with a Postgraduate Degree – Raw coefficients*

Outcome	(1) Sentiment	(2) Rating
1(Booker=1)×Post	-0.0121***	-0.124***
	(0.00324)	(0.0148)
1(Booker=1)×Post×Share of Judges with a Postgraduate Degree	-0.0447*	-0.0634
	(0.0264)	(0.108)
Observations	6,681,756	6,681,756

Notes: The unit of observation is a review. The dependent variable is reported at the top of each column. *Sentiment* refers to the sentiment valence of the review (negative or positive) and *Rating* to its star rating (number of stars). The model specification follows Equation 8 where the variable *Share of Judges with a Postgraduate Degree* and its interactions with $1(Booker = 1) \times Post$ and *Post* are included (see footnote 12 in the main text). The term $1(Booker = 1)$ is an indicator that takes value one if a book is awarded the Booker, and *Post* is an indicator for the post-Booker period. Accordingly, $1(Booker = 1) \times Post$ measures the difference in sentiment and rating between awarded and non-awarded books, conditional on controls. The variable *Share of Judges with a Postgraduate Degree* refers to the share of judges with a postgraduate degree. Each specification includes book fixed effects, a flexible control for the number of days that elapsed between the publication of the review and the publication of the book, and a flexible control for the number of days that elapsed between the publication of the review and the attribution of the Booker. Standard errors clustered at the book title level are reported in parentheses. ***Significant at 1% level; **significant at 5% level; *significant at 10% level.

D Survey Evidence

Figure D.1a: What makes you want to buy a particular book? - The literary prize(s) it has received.

Figure D.1b: Please indicate the extent to which you agree with the following statements - If I hesitate between two books, I am more likely to buy the one which has received a literary prize.

Figure D.1c: Please indicate the extent to which you agree with the following statements. - I usually expect the awarded books I read to be better than what they actually are.

Figure D.1: *Additional Outcomes of the Online Survey on Reading Habits*

Figure D.1a

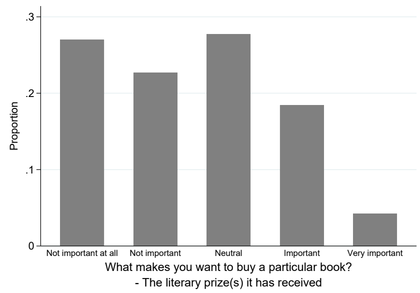


Figure D.1b

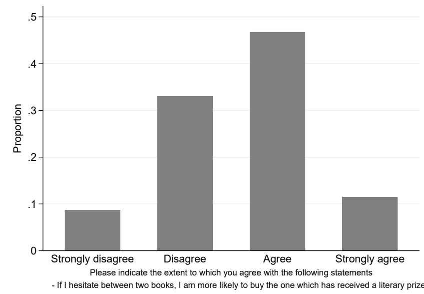
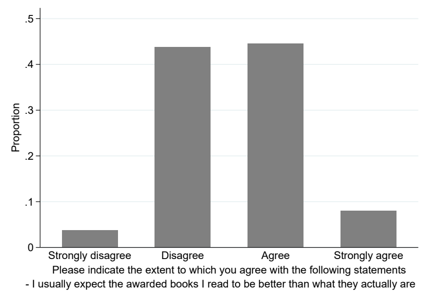


Figure D.1c



Notes: The unit of analysis is a survey respondent.

E Derivation of the Welfare Formula

Let $\widetilde{U}_{ij}$ be the utility consumer i expects to get from consuming book j (called "decision utility") and U_{ij} the utility consumer i actually obtains from consuming book j (called "experienced utility"). The difference between experienced and decision utility is then given by d_{ij}, such that

$$d_{ij} = U_{ij} - \widetilde{U}_{ij}.$$

When $d_{ij} > 0$, book j is better than what the consumer expected; when $d_{ij} < 0$, the book is worse.

Rational consumers maximize their decision utility $\widetilde{U}_{ij}$ but receive utility U_{ij}. Following Train [2015], let us assume that the book that gives the consumer the highest decision utility is $j*$ and the book that gives her the highest experienced utility is $k*$. If consumers have imperfect knowledge and overestimate the utility they will receive from reading a book, then $j* \neq k*$ and $d_{ij} < 0$. The utility loss borne by the consumer is thus $U_{j*} - U_{k*}$.

As Train [2015] shows, the average consumer surplus (CS) can be expressed as

$$CS = \frac{1}{\alpha}E(U_{j*}) = \frac{1}{\alpha}E(\widetilde{U}_{j*} + d_{j*})$$

and the average loss in surplus due to imperfect knowledge is given by

$$\Delta CS = \frac{1}{\alpha}E(U_{j*}) - \frac{1}{\alpha}E(U_{k*}) = \frac{1}{\alpha}E(\widetilde{U}_{j*} + d_{j*}) - \frac{1}{\alpha}E(U_{k*}),$$

where $E(\widetilde{U}_{j*})$ is consumer's expectation of the maximum value of her decision utility and $E(d_{j*})$ is the average difference between experienced and decision utility (Train [2015]).

Absent the Booker, the loss in surplus is

$$\Delta CS^c = \frac{1}{\alpha}E(U_{j^c*}) - \frac{1}{\alpha}E(U_{k*}).$$

The change in consumer surplus from the status quo where the Booker exists to the counterfactual scenario absent the Booker is therefore given by

$$\begin{aligned}\Delta CS - \Delta CS^c &= \frac{1}{\alpha}E(U_{j*}) - \frac{1}{\alpha}E(U_{j^c*}) \\ &= \frac{1}{\alpha}\Big[E(\widetilde{U}_{j*} - \widetilde{U}_{j^c*}) + E(d_{j*} - d_{j^c*})\Big].\end{aligned}$$

Given the modelling assumptions of our nested logit model (Train [2009, 2015]):

- $E(\widetilde{U}_{j*} - \widetilde{U}_{j^c*}) = \ln\left(1 + \left[\sum_j \exp\left(\frac{\tilde{\delta}_j}{1-\sigma}\right)\right]^{1-\sigma}\right) - \ln\left(1 + \left[\sum_j \exp\left(\frac{\tilde{\delta}^c_j}{1-\sigma}\right)\right]^{1-\sigma}\right)$, where $\tilde{\delta}_j$ and $\tilde{\delta}^c_j$ are the mean utility consumers expect to get from consuming book j in the status quo and in the counterfactual, respectively. The first part of the expression is consumer expected surplus in the status quo, and the second part is consumer expected surplus in the counterfactual – that is, absent the Booker.

- $E(d_{j*}) = E(U_{j*} - \widetilde{U}_{j*}) = \sum_j s_j(\delta_j - \widetilde{\delta}_j)$, where s_j is book j's market share in the status quo and δ_j is the mean utility consumers actually obtain from consuming book j. The expression reflects the fact that, in the status quo, consumers take decisions based on $\widetilde{\delta}_j$ (decision utility) but obtain δ_j (experienced utility).

- $E(d_{j^c*}) = E(U_{j*} - \widetilde{U}_{j^c*}) = \sum_j s^c_j(\delta_j - \widetilde{\delta}^c_j)$, where s^c_j is book j's market share absent the Booker. The expression reflects the fact that, in the counterfactual, consumers take decisions based on $\widetilde{\delta}_j^c$ (decision utility absent the Booker) but obtain a mean utility equal to the mean utility under the status quo δ_j.

Putting everything together, we have the average change in consumer surplus is given by the following formula:

$$\Delta CS = \frac{1}{\alpha}\Bigg\{ \ln\left(1 + \left[\sum_j \exp\left(\frac{\tilde{\delta}_j}{1-\sigma}\right)\right]^{1-\sigma}\right) - \ln\left(1 + \left[\sum_j \exp\left(\frac{\tilde{\delta}_j^c}{1-\sigma}\right)\right]^{1-\sigma}\right)$$
$$+ \sum_j s_j(\delta_j - \tilde{\delta}_j) - \sum_j s_j^c(\delta_j - \tilde{\delta}_j^c)\Bigg\}.$$

F Estimating the Substitution Parameter σ

To estimate σ, we leverage our sales data from the bestseller lists published by Publishers Weekly, which contains weekly sales for 7,379 editions. Then, as in Berry [1994], we estimate σ by running the following regression:

$$\ln(s_{jt}) - \ln(s_{0t}) = \sigma \ln\left(\frac{s_{jt}}{1-s_{0t}}\right) + \mu_j,$$

where $s_{jt} = q_{jt}/M$, $s_{0t} = 1 - Q_t/M$, and μ_j are edition fixed effects. The term q_{jt} refers to the sales of book j in week t, Q_t to the total book sales in week t (based on the titles in our dataset), and M to the market size. Because we assume that each American is making a discrete decision every month between buying a book or consuming the outside good and our data are at the weekly level, the market size is equal to M = US population size*0.25.

Since $\ln\left(\frac{s_{jt}}{1-s_{0t}}\right)$ is by construction endogenous, we need an instrument to consistently estimate σ. Usually, one would use the total number of titles available each week as instrument. By construction, this number in our data is always the same. Therefore, we instead use the number of titles available each month, exploiting the variations caused by the fact that the same book can be a weekly bestseller during several different weeks in the same month. We obtain $\sigma = 0.376$.

References

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2022). Learning From Reviews: The Selection Effect and the Speed of Learning. *Econometrica*, 90(6), 2857–2899. [pp. 6]

Adler, M. (1985). Stardom and Talent. *The American Economic Review*, 75(1), 208–212. [pp. 9, 37]

Aguiar, L. and Waldfogel, J. (2018). Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music. *Journal of Political Economy*, 126(2), 492–524. [pp. 30, 33]

Aguiar, L. and Waldfogel, J. (2019). Digitization and the welfare effects of motion picture trade. [pp. 33]

Akbik, A., Bergmann, T., Blythe, D., Rasul, K., Schweter, S., and Vollgraf, R. (2019). FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)* (pp. 54–59). Minneapolis, Minnesota: Association for Computational Linguistics. [pp. 11]

Allcott, H. (2011). Consumers' Perceptions and Misperceptions of Energy Costs. *American Economic Review*, 101(3), 98–104. [pp. 5, 7]

Allcott, H. (2013). The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market. *American Economic Journal: Economic Policy*, 5(3), 30–66. [pp. 5, 31]

Armstrong, T. B. and Kolesár, M. (2018). Optimal Inference in a Class of Regression Models. *Econometrica*, 86(2), 655–683. [pp. 40, 41]

Ashenfelter, O. and Jones, G. V. (2013). The Demand for Expert Opinion: Bordeaux Wine. *Journal of Wine Economics*, 8(3), 285–293. [pp. 2]

Ashworth, J., Heyndels, B., and Werck, K. (2010). Expert judgements and the demand for novels in Flanders. *Journal of Cultural Economics*, 34(3), 197–218. [pp. 6]

Asmat, R., Borowiecki, K. J., and Law, M. T. (2023). Do experts and laypersons differ? Some evidence from international classical music competitions. *Journal of Economic Behavior Organization*, 214, 270–290. [pp. 3]

Bartoshuk, L. (2014). The Measurement of Pleasure and Pain. *Perspectives on Psychological Science*, 9(1), 91–93. PMID: 26173247. [pp. 2]

Bell, R. M. and McCaffrey, D. F. (2002). Bias reduction in standard errors for linear regression with multi-stage samples. *Survey Methodology*, 28(2), 169–182. [pp. 22, 44, 45]

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics*, 25(2), 242–262. [pp. 30, 32, 33, 34, 53]

Berry, S. T. and Haile, P. A. (2021). Foundations of demand estimation. In *Handbook of Industrial Organization*, volume 4 of *Handbook of Industrial Organization* (pp. 1–62). Elsevier. [pp. 33]

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119(1), 249–275. [pp. 22]

Bourdieu, P. (1979). *La distinction. Critique sociale du jugement.* Paris, éd. de Minuit. [pp. 2]

Bourdieu, P. (1983). The field of cultural production, or: The economic world reversed. *Poetics*, 12(4), 311–356. [pp. 2, 37]

Brynjolfsson, E., Hu, Y. J., and Smith, M. D. (2003). Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science*, 49(11), 1580–1596. [pp. 11, 16]

Butler, M., Hausmann, A., and Kirchhofer, A. (2016). *Precarious Alliances. Cultures of Participation in Print and Other Media.* transcript Verlag. [pp. 9]

Cameron, S. (1995). On the role of critics in the culture industry. *Journal of Cultural Economics*, 19(4), 321–331. [pp. 2, 6]

Cardell, N. S. (1997). Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity. *Econometric Theory*, 13(2), 185–213. [pp. 31]

Chevalier, J. and Goolsbee, A. (2003). Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com. *Quantitative Marketing and Economics*, 1(2), 203–222. [pp. 11, 16]

D'agostino, R. B., Belanger, A., and Jr., R. B. D. (1990). A Suggestion for Using Powerful and Informative Tests of Normality. *The American Statistician*, 44(4), 316–321. [pp. 43]

Deuchert, E., Adjamah, K., and Pauly, F. (2005). For Oscar Glory Or Oscar Money? *Journal of Cultural Economics*, 29(3), 159–176. [pp. 6]

Dobrescu, L. I., Luca, M., and Motta, A. (2013). What makes a critic tick? Connected authors and the determinants of book reviews. *Journal of Economic Behavior Organization*, 96, 85–103. [pp. 2]

English, J. F. (2014). Chapter 6 - The Economics of Cultural Awards. In V. A. Ginsburgh andd D. Throsby (Eds.), *Handbook of the Economics of Art and Culture*, volume 2 of *Handbook of the Economics of Art and Culture* (pp. 119–143). Elsevier. [pp. 2, 3]

Friberg, R. and Grönqvist, E. (2012). Do Expert Reviews Affect the Demand for Wine? *American Economic Journal: Applied Economics*, 4(1), 193–211. [pp. 6]

Gelman, A. and Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3), 447–456. [pp. 40]

Ginsburgh, V. (2003). Awards, Success and Aesthetic Quality in the Arts. *Journal of Economic Perspectives*, 17(2), 99–111. [pp. 2, 6]

Ginsburgh, V. (2016). *On Judging Art and Wine*, (pp. 245–265). Springer International Publishing: Cham. [pp. 3]

Ginsburgh, V., Radermecker, A.-S., and Tommasi, D. (2019). The effect of experts' opinion on prices of art works: The case of Peter Brueghel the Younger. *Journal of Economic Behavior Organization*, 159, 36–50. [pp. 6]

Ginsburgh, V. and Weyers, S. (2014). Nominees, winners, and losers. *Journal of Cultural Economics*, 38(4), 291–313. [pp. 6]

Ginsburgh, V. A. and van Ours, J. C. (2003). Expert Opinion and Compensation: Evidence from a Musical Competition. *American Economic Review*, 93(1), 289–296. [pp. 6]

Haan, M. A., Dijkstra, S. G., and Dijkstra, P. T. (2005). Expert Judgment Versus Public Opinion – Evidence from the Eurovision Song Contest. [pp. 3]

Hausman, C. and Rapson, D. S. (2018). Regression Discontinuity in Time: Considerations for Empirical Applications. *Annual Review of Resource Economics*, 10(1), 533–552. [pp. 21, 40, 41]

Hilger, J., Rafert, G., and Villas-Boas, S. (2011). Expert Opinion and the Demand for Experience Goods: An Experimental Approach in the Retail Wine Market. *The Review of Economics and Statistics*, 93(4), 1289–1296. [pp. 6]

Holbrook, M. B. (1999). Popular Appeal versus Expert Judgments of Motion Pictures. *Journal of Consumer Research*, 26(2), 144–155. [pp. 3]

Hutto, C. and Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. [pp. 46]

Hörner, J. and Lambert, N. S. (2021). Motivational Ratings. *The Review of Economic Studies*, 88(4), 1892–1935. [pp. 6]

Imbens, G. W. and Kolesár, M. (2016). Robust Standard Errors in Small Samples: Some Practical Advice. *The Review of Economics and Statistics*, 98(4), 701–712. [pp. 22, 44, 45]

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), 615–635. The regression discontinuity design: Theory and applications. [pp. 42]

Jin, G. Z. and Sorensen, A. T. (2006). Information and consumer choice: The value of publicized health plan ratings. *Journal of Health Economics*, 25(2), 248–275. [pp. 7]

Kahneman, D. (1994). New Challenges to the Rationality Assumption. *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft*, 150(1), 18–36. [pp. 5, 7, 30, 31]

Kolesár, M. and Rothe, C. (2018). Inference in Regression Discontinuity Designs with a Discrete Running Variable. *American Economic Review*, 108(8), 2277–2304. [pp. 40, 41]

Lagios, N. and Méon, P.-G. (2023). Experts, Information, Reviews, and Coordination: Evidence on How Prizes Affect Sales. *Journal of Industrial Economics*, forthcoming. [pp. 3, 6, 24, 37]

Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2), 655–674. The regression discontinuity design: Theory and applications. [pp. 40]

Leibenstein, H. (1950). Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand. *The Quarterly Journal of Economics*, 64(2), 183–207. [pp. 23]

Lewbel, A. (1997). Constructing Instruments for Regressions With Measurement Error When no Additional Data are Available, with An Application to Patents and RD. *Econometrica*, 65(5), 1201–1213. [pp. 21, 43, 44]

Lewbel, A. (2012). Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models. *Journal of Business & Economic Statistics*, 30(1), 67–80. [pp. 21, 42, 43]

Lien, A. K., Randem, L. M., Taralrud, H. P. F., and Edalati, M. (2022). OSN Dashboard Tool For Sentiment Analysis. [pp. 11]

Loeper, A., Steiner, J., and Stewart, C. (2014). Influential Opinion Leaders. *The Economic Journal*, 124(581), 1147–1167. [pp. 6, 37]

Loria, S. *et al.* (2018). textblob Documentation. *Release 0.15*, 2(8), 269. [pp. 46]

Mahrukh, R., Shakil, S., and Malik, A. S. (2023). Sentiments analysis of fMRI using automatically generated stimuli labels under naturalistic paradigm. *Scientific Reports*, 13(1), 7267. [pp. 11]

Moseley, M. *et al.* (2019). How the Booker Prize Won the Prize. *American, British and Canadian Studies*, (33), 206–221. [pp. 9]

Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy*, 78(2), 311–329. [pp. 7]

Nevo, A. (2000). A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand. *Journal of Economics & Management Strategy*, 9(4), 513–548. [pp. 33]

Pang, B. and Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135. [pp. 4, 11]

Ponzo, M. and Scoppa, V. (2015). Experts' awards and economic success: evidence from an Italian literary prize. *Journal of Cultural Economics*, 39(4), 341–367. [pp. 3, 6]

Reimers, I. (2019). Copyright and Generic Entry in Book Publishing. *American Economic Journal: Microeconomics*, 11(3), 257–84. [pp. 14, 33]

Reimers, I. and Waldfogel, J. (2017). Throwing the Books at Them: Amazon's Puzzling Long Run Pricing Strategy. *Southern Economic Journal*, 83(4), 869–885. [pp. 18]

Reimers, I. and Waldfogel, J. (2021). Digitization and Pre-purchase Information: The Causal and Welfare Impacts of Reviews and Crowd Ratings. *American Economic Review*, 111(6), 1944–71. [pp. 4, 6, 7, 13, 18, 30, 33, 34]

Reinstein, D. A. and Snyder, C. M. (2005). The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics. *The Journal of Industrial Economics*, 53(1), 27–51. [pp. 6]

Robbins, L. (1938). Interpersonal Comparisons of Utility: A Comment. *The Economic Journal*, 48(192), 635–641. [pp. 2]

Rogge, E., Nevens, F., and Gulinck, H. (2007). Perception of rural landscapes in Flanders: Looking beyond aesthetics. *Landscape and Urban Planning*, 82(4), 159–174. [pp. 3]

Roodman, D., Ørregaard Nielsen, M., MacKinnon, J. G., and Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal*, 19(1), 4–60. [pp. 22, 44, 45]

Rossi, M. (2021). Quality Disclosures and Disappointment: Evidence from the Academy Awards. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, EC '21 (pp. 790–791). New York, NY, USA: Association for Computing Machinery. [pp. 6, 23]

Train, K. (2015). Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples. *Journal of Choice Modelling*, 16, 15–22. [pp. 7, 30, 51, 52]

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press. [pp. 52]

Villanes, A. and Healey, C. G. (2023). Domain-specific text dictionaries for text analytics. *International Journal of Data Science and Analytics*, 15(1), 105–118. [pp. 11]