Effect of Literary Awards on Reader Response

Jungeun "June" Lim | March 29, 2020

Abstract

How does winning a literary award affect the readers' evaluation of the book? Previous studies have identified an ironic effect of receiving an award; while award-winners receive more reviews, the reviews they get become more negative towards them. However, whether such a pattern can be generalized to books that are not written in English but translated to English has not been studied. This report shows that, while this phenomenon also occurs among books that are translated into English, the degree of effect could be different. In particular, it finds that 1) the effect of winning an award on the number of reviews tends to be bigger, and 2) the effect of winning an award on the review ratings tends to be smaller for translated books compared to those for books originally written in English.

Introduction

Awards and prizes affect the recipients and the market in many different and interesting ways. They obviously boost the reputation of the recipients and increase the sales of the products created by or related to them (Ponzo & Scoppa, 2015); how strong the effect of an award is and how long it lasts have been studied in various academic fields (Frey & Gallus, 2017). However, there are other potential consequences of awards that are not as obvious as the immediate positive effect. For instance, Kovács and Sharkey (2014) studies four well-known literary awards for books written in English and shows that the public's evaluation of the award winner (measured as the average of star ratings by online reviewers) become worse on average after the announcement of the award. They also suggest one of the reasons behind the phenomenon; the award attracts people who would not have attempted to read the book if it did not receive the award, and they tend to get disappointed after reading it.

Although this argument is persuasive, whether it can be generalized to books that are not written in English but translated to English has not been studied. Because of the different nature of translated books and their readers, some of the mechanisms behind the effect of awards can act differently. Translated literature makes up just two to three percent of literature published in English-speaking countries; even though translated books and imported cultural products in general are being consumed more widely in recent years in English-speaking countries, people who frequently read translated books could still be a selective group that is different from the general public. For instance, they may have a clear intention to read a particular book from a particular country or have been looking for something different from the books written in English by English-speaking authors from the first place. Based on this motivation, this report aims to identify the effect of awards on translated books and compare it to the effect of awards on books originally written in English.

Data

I investigated the difference between the mere nominees (short-listed books that did not receive the award) and winners of the Booker Prize and the International Booker Prize. The Booker Prize is given to a book originally written in English, and the International Booker Prize is given to a book translated from any non-English languages to English. Since both awards carry the same name, I decided that the winners and nominees of the Booker Prize and International Booker Prize could be a good proxy to be used to answer my question.

I limited my focus to a four-year period for consistency (from 2016 to 2019, inclusive). Whereas the Booker Prize has been given to a book, the International Booker Prize was previously given to an author rather than a book before 2016. Since 2016, both awards are given annually to a book.

My dataset includes 1) the number of reviews and 2) the star ratings of the reviews written by the users of Goodreads.com. 6886 reviews of 23 books originally written in English and nominated for the Booker Prize and 5300 reviews of 24 translated books that were nominated for the International Booker Prize were scraped using R scripts. Since Goodreads.com displays up to 300 reviews of a book, I could scrape only up to 300 most helpful English reviews of each book. If a book had received less than 300 reviews, I could scrape all the reviews of it. However, the majority of the books except for a few nominees for the International Booker Prize had more than 300 reviews, which should be taken into consideration when interpreting the result.

Figure 1 to 4 show the differences in the number of reviews and star ratings that the award winner and mere nominees received before and after the announcement of the award.

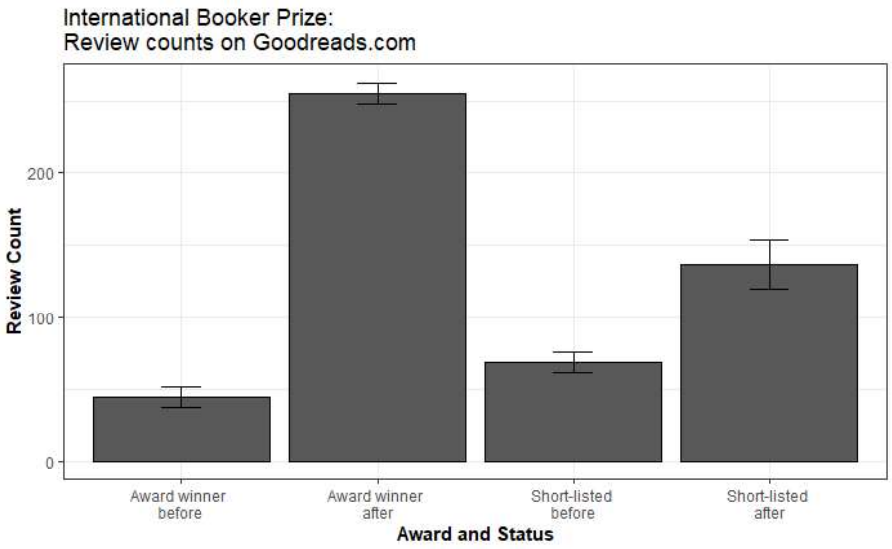

Figure 1. Mean and standard error of the number of reviews on Goodreads.com for award winners and short-listed books before and after the award announcement (International Booker Prize).

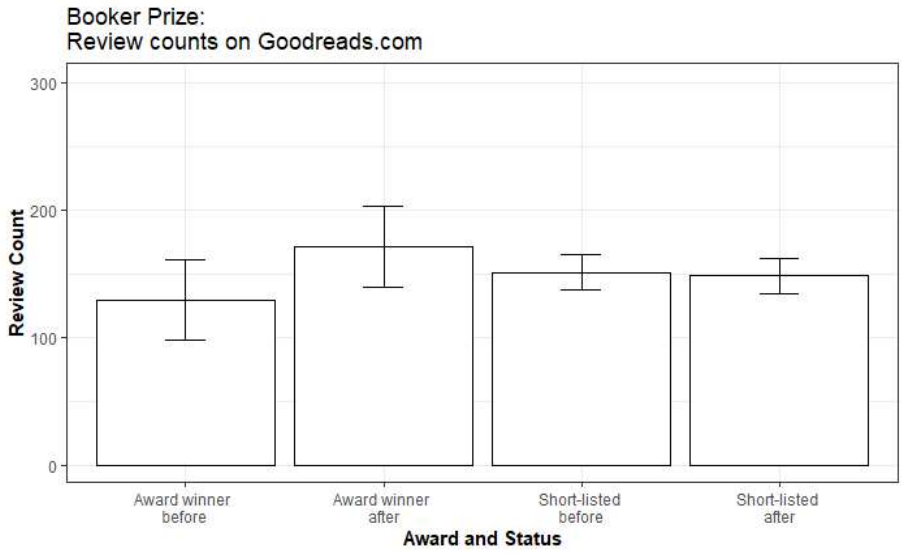

Figure 2. Mean and standard error of the number of reviews on Goodreads.com for award winners and short-listed books before and after the award announcement (Booker Prize).

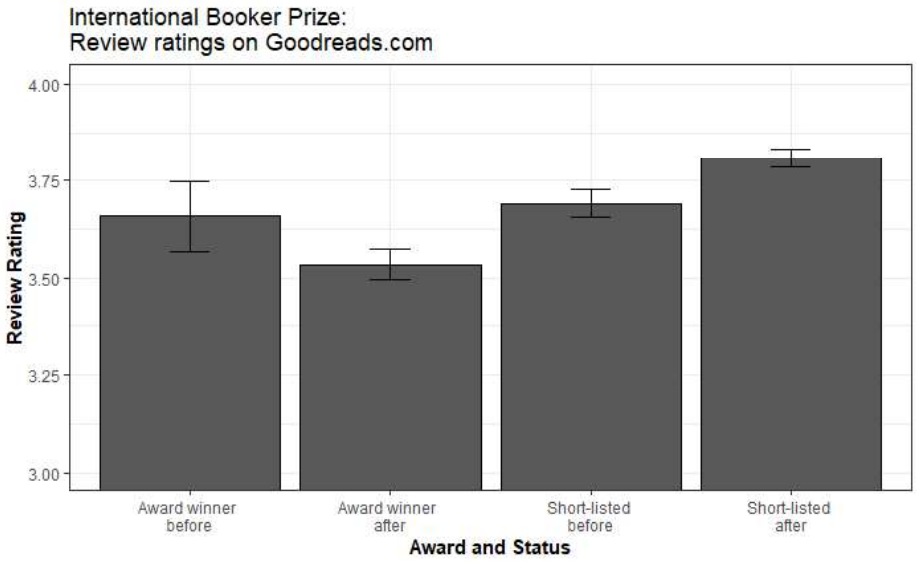

Figure 3. Mean and standard error of star ratings on Goodreads.com for award winners and short-listed books before and after the award announcement (International Booker Prize).

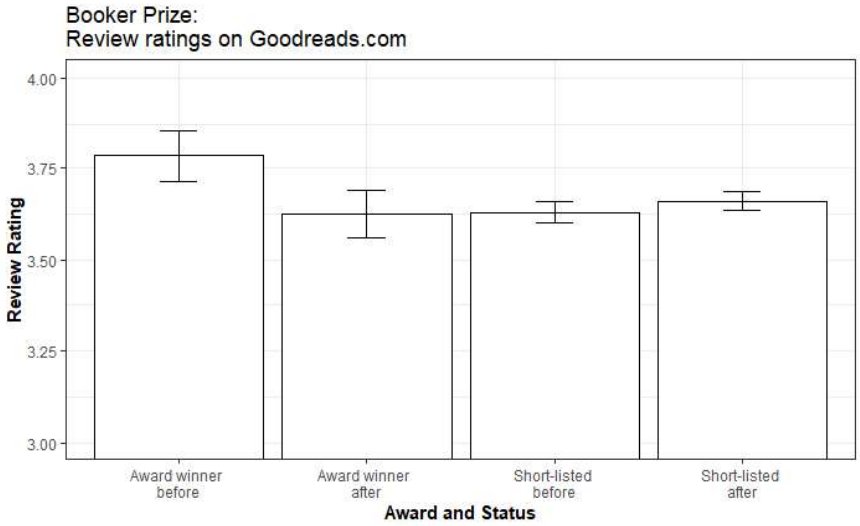

Figure 4. Mean and standard error of star ratings on Goodreads.com for award winners and short-listed books before and after the award announcement (Booker Prize).

To construct Figure 1 and 3, I counted the number of reviews that each book has received before and after the announcement of the award. To construct Figure 2 and 4, I grouped all the reviews into the "before" group and "after" group and computed the mean star ratings.

We can see from Figure 1 and 2 that a) the award winners get more reviews after the award was announced than the mere nominees in general and b) this trend seems to be stronger for translated books. We can see from Figure 3 and 4 that a) the award winners get more negative reviews after the announcement of the award while the mere nominees get more positive reviews after the announcement of the award in general and b) this trend seems to be weaker for translated books. The statistical details of these observations will be discussed later in this report.

Method: Difference-in-Difference

Difference-in-difference is a widely used econometrics technique that is used to approximate experiments from non-experimental data. To use difference-in-difference, you need a panel dataset and should be able to find a reasonable control group and a treatment group from it. Since you know the underlying difference between the control group and treatment group when neither was treated (before the treatment), you could estimate the unbiased effect of the treatment.

Many prior studies on the effect of awards have used this approach to make causal estimates utilizing the nominees as a control group and winners as a treatment group. Even though winning an award is not a random treatment and the literary quality is impossible to be objectively and numerically measured, it is still a reasonable assumption that the nominees and the winners are of similar quality. I followed this convention in this assignment.

Translated Books (Booker International Prize)	
	Dependent variable:
	review_count
winner	-23.950
	(29.816)
after	68.100***
	(17.214)
winner:after	142.900***
	(42.166)
Constant	68.450***
	(12.172)
Observations	48
R2	0.536
Adjusted R2	0.504
Residual Std. Error	54.436 (df = 44)
F Statistic	16.925*** (df = 3; 44)
Note:	*p<0.1; **p<0.05; ***p<0.01

Books Written in English (Booker Prize)	
	Dependent variable:
	review_count
winner	-21.889
	(30.459)
after	-2.556
	(20.084)
winner:after	44.556
	(43.076)
Constant	150.889***
	(14.202)
Observations	46
R2	0.028
Adjusted R2	-0.041
Residual Std. Error	60.253 (df = 42)
F Statistic	0.410 (df = 3; 42)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 1. The effect of winning the International Booker Prize from an OLS regression model that predicts the number of reviews that a book receives on Goodreads.com. Left: books translated to English (nominees and winners for the International Booker Prize), Right: books written in English (nominees and winners for the Booker Prize)

Effect of Winning Award on the Number of Reviews and the Review Ratings

Table 1 shows the result of the OLS regression conducted with two categorical independent variables ("before or after the announcement of the award" and "winner or mere nominee") to predict the number of reviews. While we can see winning the award increases the number of reviews that the book receives in general, the effect is stronger and more significant for translated books.

Table 2 shows the result of the OLS regression conducted with two categorical independent variables ("before or after the announcement of the award" and "winner or mere nominee") to predict the review ratings. While we can see winning the award decreases the star ratings that the book receives in general, the effect is weaker and less significant for translated books.

Conclusion, Limitation and Future Work

The OLS regression result suggests that award-winning books tend to attract more readers after the announcement of the award in general and this effect is more significant for translated books. With what I have now, I cannot say anything about the mechanisms behind these differences, but maybe it is because translated books are not advertised or reviewed on mass media much. Therefore, if they win awards they may suddenly get increased attention. On the other hand, books written in English may already have some publicity before winning the award.

Translated Books (Booker International Prize)

	Dependent variable:
	review_rating
winner	-0.012
	(0.107)
after	0.129***
	(0.046)
winner:after	-0.188
	(0.119)
Constant	3.690***
	(0.038)
Observations	4,230
R2	0.004
Adjusted R2	0.004
Residual Std. Error	1.221 (df = 4226)
F Statistic	6.318*** (df = 3; 4226)
Note:	*p<0.1; **p<0.05; ***p<0.01

Books Written in English (Booker Prize)

	Dependent variable:
	review_rating
winner	0.155**
	(0.073)
after	0.031
	(0.040)
winner:after	-0.189**
	(0.096)
Constant	3.629***
	(0.030)
Observations	5,086
R2	0.001
Adjusted R2	0.0004
Residual Std. Error	1.293 (df = 5082)
F Statistic	1.607 (df = 3; 5082)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 2. The effect of winning the International Booker Prize from an OLS regression model that predicts the star ratings of reviews that a book receives on Goodreads.com. Left: books translated to English (nominees and winners for the International Booker Prize), Right: books written in English (nominees and winners for the Booker Prize)

The OLS regression result also suggests that award-winning books tend to receive more unfavourable reviews after the announcement of the award, and this effect seems less significant for translated books. It could be because the readers of translated books are more ready for something different and unexpected, and as a result, give more generous ratings to translated books.

There are some critical limitations in this analysis, mostly in terms of the amount of the data analyzed. First, I only analyzed the four-years worth of data. Second, Goodreads.com has much more reviews but I could not analyze all of them because Goodreads.com shows only up to 300 reviews. As a result, my dataset was noisy and may not be representative enough of the readers' responses to the books. Future work could analyze more data, perhaps the reviews of the same books from Amazon.com and Amazon.co.uk for a more robust analysis.

Supplementary Material

R and Python code to scrape, preprocess, analyze, and visualize the data is available in the following GitHub link: https://github.com/JuneJLim/INF2178_a4

Ethical Issues and Concerns: Issues of Using Online Reviews for Research

1-1. Lack of Consent from the Authors and the Website

While web scraping for research in academia has become a common practice, whether it is ethical to collect and analyze data from online without consent from the authors remains unclear. It would have been unrealistic for me to acquire informed consent from everyone who posted a review on Goodreads.com that I collected for this assignment; however, it does not mean that I can blindly justify using these data because I could access them freely.

The terms of use of the website should also be considered when scraping the data from it. While Goodreads.com grants "a limited, non-exclusive, non-transferable, non-sublicensable license to access and make personal and non-commercial use of the Service" to the users, it also specifies that "any collection and use of any book listings, descriptions, reviews or other material included in the Service" and "any use of data mining, robots, or similar data gathering and extraction tools" are not included in this license and the rights for doing such are "reserved and retained by us or our licensors, suppliers, publishers, rightsholders, or other content providers."[1] While students who do class projects like this report and even academic researchers in general often scrape websites without explicitly discussing it with the companies that own the websites, they should be aware of the possible legal issues and consequences of doing so.

1-2. Potential Harm to Privacy

Although republishing the data along with the code used to analyze them is necessary to ensure the replicability of the research, doing so could potentially lead to compromising the privacy of the individuals who created the data or were involved with them. Anonymizing data is not always a solution because it could be possible to deanonymize the data by linking them with other datasets.

1-3. Overloading the Website

The server of a website could get overloaded if it gets too many requests in a short period of time. While Goodreads.com is a large website that could handle high traffic, I put a delay between each request to ensure that I was not overloading the server and to avoid being blocked.

[1] https://www.goodreads.com/about/terms

References (Academic Papers)

Dai, H., Dietvorst, B. J., Tuckfield, B., Milkman, K. L., & Schweitzer, M. E. (2017). Quitting When the Going Gets Tough: A Downside of High Performance Expectations. *Academy of Management Journal*, *61*(5), 1667–1691. https://doi.org/10.5465/amj.2014.1045

Dekker, E., & de Jong, M. (2018). What Do Book Awards Signal? An Analysis of Book Awards in Three Countries. *Empirical Studies of the Arts*, *36*(1), 3–21. https://doi.org/10.1177/0276237416689636

Jensen, M., & Kim, H. (2015). The Real Oscar Curse: The Negative Consequences of Positive Status Shifts. *Articles and Chapters*. https://scholarship.sha.cornell.edu/articles/1172

Frey, B. S., & Gallus, J. (2017). Towards an Economics of Awards. *Journal of Economic Surveys*, *31*(1), 190–200. https://doi.org/10.1111/joes.12127

Kovács, B., & Sharkey, A. J. (2014). The Paradox of Publicity: How Awards Can Negatively Affect the Evaluation of Quality. *Administrative Science Quarterly*. https://doi.org/10.1177/0001839214523602

Ponzo, M., & Scoppa, V. (2015). Experts' awards and economic success: Evidence from an Italian literary prize. *Journal of Cultural Economics*, *39*(4), 341–367. https://doi.org/10.1007/s10824-015-9239-4

Reschke, B. P., Azoulay, P., & Stuart, T. E. (2018). Status Spillovers: The Effect of Status-conferring Prizes on the Allocation of Attention. *Administrative Science Quarterly*, *63*(4), 819–847. https://doi.org/10.1177/0001839217731997

Sasaki, S., Kurokawa, H., & Ohtake, F. (2019). Positive and negative effects of social status on longevity: Evidence from two literary prizes in Japan. *Journal of the Japanese and International Economies*, *53*, 101037. https://doi.org/10.1016/j.jjie.2019.101037

References (Data and Software)

Hlavac, M. (2018). *stargazer: Well-Formatted Regression and Summary Statistics Tables*. R package version 5.2.2. https://CRAN.R-project.org/package=stargazer

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Wickham, H., Romain, F., Henry, L., and Müller, K. (2018). *dplyr: A Grammar of Data Manipulation*. R package version 0.7.6. https://CRAN.R-project.org/package=dplyr

Wickham., H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer Verlag.