I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15501

Famous after Death: The Effect of a Writer's Death on Book Sales

Michela Ponzo
Vincenzo Scoppa

AUGUST 2022

DISCUSSION PAPER SERIES

IZA DP No. 15501

Famous after Death: The Effect of a Writer's Death on Book Sales

Michela Ponzo
University of Calabria

Vincenzo Scoppa
University of Calabria and IZA

AUGUST 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Famous after Death: The Effect of a Writer's Death on Book Sales*

In the standard neoclassical model consumers use all the available information and the demand for goods depends exclusively on preferences and prices whereas other spurious information do not play any role. In the market for books, we investigate if – in contrast to the standard model – the death of a writer has an impact on demand for his/her books, that is, we ask if consumers are affected by factors such as emotions and limited attention, as highlighted in behavioral economics. We use bestseller lists at week level for about 30 years (1975-2005) and through a Regression Discontinuity Design we evaluate the impact of a writer's death on the probability of entering in the bestseller list in the period immediately following his/her death. Controlling for age, gender, literary prizes, publishers' relevance and time dummies we find that a writer's death increases the probability of being in the bestseller list of more than 100%. Using a non-parametric RD approach we find very similar results. A number of robustness checks – changing the time window around the death, the estimation method, the outcome variable, the sample used – confirm our findings. In the attempt to investigate which mechanism drives consumers' decisions, we find a much greater impact for writers dying at an early age, for more famous writers and when the news is covered more extensively, suggesting that emotions and media attention are the main drivers of the impact.

JEL Classification: D91, Z10, Z11, L82, M30, D12

Keywords: book sales, writer's death, emotions, limited attention, salience, cultural economics, behavioral economics

Corresponding author:
Vincenzo Scoppa
Dipartimento di Economia, Statistica e Finanza
Università della Calabria
Via Ponte Bucci
87036 Arcavacata di Rende (CS)
Italy
E-mail: v.scoppa@unical.it

* We would like to thank Concetta Castiglione, Vincenzo Carrieri, Maria De Paola and Matteo Marini for useful comments and suggestions.

1. Introduction

Umberto Eco, a famous Italian medievalist, novelist and semiotician died on 19th February 2016. His most famous novel, "The Name of the Rose", published in 1981, after about 35 years from its first publication, was in the sixth position in the bestseller list in the following week and remained in the bestseller list for several weeks. Other Umberto Eco's five books entered in the top 20 bestseller list. Italo Calvino, another famous Italian novelist, died on 19th September 1985. The following week six of his novels (published several years earlier) entered in the top 20 bestseller list and remained among bestsellers for weeks.

In the traditional economic model, the demand for goods depends exclusively on consumer preferences, prices and income; furthermore, the consumers are assumed to have perfect information and to decide on the basis of these: therefore, other spurious information are predicted to have no effect on consumers' behavior.

In this paper we deal with the demand for books and we investigate how an exogenous event as the death of a writer affect consumers' decisions. According to the standard model, demand for books – like other products – should not depend on an event as the death of a writer, which typically does not change preferences, does not communicate new information on the quality of books, and so on.[1] The death of an author can be considered one of the "Seemingly Irrelevant Factors" (Thaler, 2016) – a factor that should have no effect among perfectly rational individuals.

However, as shown in behavioral economics, consumers in their decisions might be affected by emotions or mood (Loewenstein, 2000; Rick and Loewenstein, 2008; Stanton et al., 2014). In contrast to the "*homo economicus*" (Thaler, 2016), "humans" get excited, sad, upset, happy and so on. These feelings can be triggered by particular events, as the death of a writer, and, as a consequence, might have an impact on individual economic decisions.

This is related to the "affect heuristic" (Slovic et al., 2002) that refers to how people make decisions that are influenced by their current emotions or the feelings toward a particular stimulus. For example, in an interesting lab experiment, Hsee and Rottenstreich (2004) show that the willingness of participants to pay for a set of CDs by Madonna is different when they are induced to do a valuation by calculation or induced to do a valuation by feelings (people's attention is steered towards emotional aspects through unconscious pre-activation). Card and Dahl (2011) show that emotional factors associated with unexpected losses by professional football teams can trigger violent episodes by men against their wives and girlfriends.

Beyond the effects of emotions, consumers can react in non-standard way to the arrival of some kind of news since typically in their decisions consumers do not use all the available information but are affected by limited attention (Simon, 1986; Gabaix, 2019; Huberman and Regev, 2001) and process only some parts of the available information while neglecting other parts. The literature has shown that the degree of inattention

[1] This might be different in the market for art since the death of an artist, by changing the future supply of his artworks, might affect the prices of his/her creations. Books can be reproduced without limits at very small marginal costs and therefore deaths should not modify supply or demand.

is affected by the salience of information and by the number of competing stimuli (DellaVigna, 2009). Therefore, although the information on an author's books, the stories and characters created, his/her writing style, etc., should be publicly available, the death of a writer is likely to make more salient the information on his/her works and stimulate consumers to buy.

In this paper – focusing on the market for books – we investigate if in the market for books consumers behave as in the traditional model or are influenced by emotions or by the salience of the information and therefore react to the news about an author's death.

We use weekly data for a period of 30 years (from 1975 to 2005) on bestsellers lists published by a leading Italian newspaper ("La Stampa"). We select a sample of writers appeared at least once in the bestseller list and complement the information regarding bestseller lists with biographical information on the date of birth and the date of death of authors. In a Regression Discontinuity Design (RDD), we evaluate the impact of the death of a writer on the probability of appearing in the bestseller list with some of his/her books in the period immediately following his/her death. Controlling for age, gender, literary awards, publisher's relevance, time dummies, we find that a writer's death increases the probability of being in the bestseller list of more than 100%.

Interestingly, corroborating our hypothesis on consumers influenced by emotions, when we separate our sample on the basis of the authors' age at death, we find a much greater impact for writers dying in an early age. The salience of information also appears to play an important role, since a larger impact occurs for already famous authors – whose death tend to receive greater attention in the news. On the other hand, the impact does not seem to depend on the relevance of the publishers.

A number of robustness checks – in which we change our outcome variable, the method of estimation, the time window we focus on before and after the author's death – strongly confirm our findings.

Our paper is related to a few other works. A number of papers have studied the impact of the death of an artist on the price of his/her paintings (Ursprung and Wiermann, 2011; Ekelund, Ressler and Watson, 2000; Maddison and Pedersen, 2008, among others).[2] The main idea is that the future supply of an artist's works will be limited after his/her death and the expectations of this relative scarcity will increase the demand for the artist's paintings.

Ekelund, Ressler and Watson (2000) examining the works of 21 Latin American artists who died between 1977 and 1996 have shown that the prices of their paintings increase significantly soon after the artists' death because of the collectors' expectations of raising prices due to the fixed supply, analogously to the Coase's (1972) conjecture for the problem of the "durable goods monopolist". Ursprung and Wiermann (2011) use a large dataset of over 400,000 transactions from art auctions to analyze if the death of an artist affects the price of his/her works. They find a hump shaped relationship between age at death and prices.

[2] Frey and Gullo (2020) analyze what happens to the reputation of economists after their death finding that the attention of other economists strongly declines for Nobel prize winners dying prematurely, while there is no effect for famous economists dying at old age.

Maddison and Pedersen (2008) using data for 93 Danish painters (died from 1983 to 2003) show that the death effect is more pronounced for artists dying prematurely since this changes much more the future supply of their works. Once they take fully into account the effect of death on future supply, they do not find any leap in prices after the news of an artist's death.

However, the fine art market analyzed in these works is a peculiar market, since an artist's death corresponds to a fixed supply of works. In other cultural markets (books, music albums, movies, sports memorabilia), goods are reproducible at a constant marginal costs (and typically are not resaled) and the described mechanism based on limited supply does not apply and there is no room for appreciation. Two papers analyzing these markets are more related to our analysis.

Matheson and Baade (2004) analyze the prices of sports trading cards of players of the Major League Baseball. They only have semi-annual data from 1990 to 2001 for 13 players. Matheson and Baade (2004) simply compare the prices of the cards 6 months before and 6 months after the death of the players, subtracting from this difference any variation in the average price of cards. The authors find that the price increases of about 11% immediately after the player's death but this higher price is not maintained in the future. The authors define the increase in price a "nostalgia spike" and attribute this to media attention surrounding a player's death.

Radford and Bloch (2013) study how the death of a celebrity affects the market for his/her memorabilia. They use data from eBay for memorabilia (autographed items and DVD movies) of 6 celebrities and compare data on prices, bids and new items offered for the 14 days before and after the celebrity's death. The authors find that the death of a celebrity generates an immediate increase in the number of items for sale and in the bid activity, followed by a subsequent reduction. These effects seem to be stronger for unexpected deaths. Less clear are the effects on prices.

In this paper we will try to verify a similar impact on the sales of books but using a much larger dataset with thousands of observations and adopting a more rigorous and modern identification strategy such as RDD.

The paper is organized as follows. Section 2 provides a brief description of the data we use. In Section 3 we conduct an econometric analysis using a parametric Regression Discontinuity approach while in Section 4 we adopt a non-parametric RD Design. In Section 5 we conduct a number of robustness checks. In Section 6 by analyzing the heterogeneity of the impact we examine the relative importance of different mechanisms at work. Section 7 offers some concluding remarks.

2. The Data

Our main dataset is based on weekly data on bestsellers lists published on "*Tuttolibri*", the cultural supplement of the leading Italian newspaper "*La Stampa*". Each week (on Saturday) *La Stampa* publishes information on bestseller list. We collected these data manually by using the digital edition of *La Stampa* freely available on the archive http://www.archiviolastampa.it/ over the period from November 8, 1975 to December 31, 2005.

We have gathered in total 1,326 weekly bestseller lists, about 44 lists per year.[3] Unfortunately, the digital archive arrives until 2005 and there are no lists available after that date.

Bestseller lists are provided by leading international data providers (currently by Nielsen BookScan, while until recently data were provided by Demoskopea Institute) on the basis of the number of copies of books sold in a representative sample of bookshops. The reference week of bestseller list is typically 14 days before the date of publication. Books are ranked separately for categories (Italian Fiction, Foreign Fiction, Non-fiction, etc.) but for our aims we only use the Italian Fiction category. The bestseller list contains from 10 to 20 titles (the number of titles in the list has changed along the sample period).

For each book we observe: the author, the title, the position in the list, the publisher, the date of publication of the bestseller list, the reference week (usually two weeks before publication) that we take into account throughout the analysis, the number of points, determined assigning 100 points to the book with the greatest number of sales in a week while the other titles in the list receive a number of points in proportion to the copies sold with respect to the first ranked book.

We transform this original data at the author level. For each author i and for each week t, starting from the first week in which he/she appeared in the bestseller list, we build the dummy *Bestseller* equal to one if in the week t the author i was present in the list with one or more books and we set *Bestseller* equal to zero if an author was not present in week t. We also build our alternative outcome variable *Points* considering the number of points as bestseller obtained in a week from an author (when more than one book from an author was present in a week we simply add the points of each book), setting to zero the points for authors not present in the bestseller list in a week.[4]

As regards biographical information, we select from Wikipedia all the individuals with a biographical entry died in each year from 1976 to 2005[5] and from the whole list we keep only entries identified as "Italian" and "Writer". We checked each entry with "*Dizionario Biografico degli Italiani*" (*Biographical Dictionary of the Italians*), that is, a biographical dictionary published by the *Enciclopedia Istituto Treccani*.[6] For each author we observe the date of death, the date of birth and the gender.

To the aim of avoiding to compare authors that are too different, in our main analysis we keep only the authors who died in the period 1976-2005, that is, we drop authors died before 1975 or died after 2005 or still alive notwithstanding they entered in the bestseller list at any time.[7]

We merged the dataset of authors with at least one presence in the bestseller list with the biographical dataset of the dead writers in the corresponding period. Our most important variable is the variable *After Death*, which is a dummy equal to one for the 6 months after the date of death of an author and zero otherwise.

[3] Typically, lists are not published in August and in some bank holidays.

[4] Two other outcome variables that we consider in the Robustness Check Section are: a) the number of books present in the bestseller list in a given week; b) an ordinal variable, *Ranking*, equal to 4 if in position 1-5; equal to 3 if in position 6-10; equal to 2 if in position 11-15; equal to 1 if in position 16-20; equal to 0 if not present in the bestseller list.

[5] For example, the entry "2005 Deaths" at the link: https://it.wikipedia.org/wiki/Morti_nel_2005

[6] The *Biographical Dictionary of the Italians* includes about 40,000 biographies of distinguished Italians. It started in 1925 and is conceived to follow the model of the Oxford Dictionary of National Biography.

[7] Alternatively, in the robustness checks we consider observations for all the writers died in the period 1976-2021. Estimating our models on this sample we find very similar results.

However, to evaluate the impact of our interest (in alternative to 6 months) we experiment considering different time windows: 1 month; 3 months, 12 months, 24 months.

For each author we observe the gender and the year of birth and we build the dummy *Female* and calculate the *Age* in each period. Finally, we also use the age at death for an analysis of the mechanisms behind the effect we analyze. Since literary prizes have been shown to have a very strong impact on book sales (Ginsburgh, 2003; for the Italian case, Ponzo and Scoppa, 2015), we collected data on the authors that obtained one of the main Italian literary prizes considering the following prizes assigned yearly: Strega; Bancarella; Campiello; Viareggio; Bagutta.[8] We build a dummy *Literary Prize* equal to one for the authors winning one of these prizes in the current or in the previous year.

Moreover, we build a dummy variable for each of the 45 publishers[9] that we use as controls in some specifications. Finally, we build the variable *Time* in weeks centered on the writer's death, that is, we set *Time* for each author equal to 0 for the week of death, increasing of one for each following week and decreasing of one for each previous week.

Table 1 reports the descriptive statistics. In our sample in total we have about 65,000 author per week observations. The authors we consider in the analysis are 111.

The probability of being in the bestseller list in a given week for an author is 5.0%. The average number of points is 1.9, the number of books 0.053. Female authors represent 21% of the sample. The average age is 67.9. About 3.4% of the authors have won some literary prizes. The average age at death is 76.4.

Table 1. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Bestseller List	65418	0.050	0.218	0	1
Points Bestseller	65418	1.907	10.743	0	224.4
Books	65418	0.053	0.241	0	6
After Death	65418	0.043	0.203	0	1
Female	65418	0.208	0.406	0	1
Age	65418	67.935	11.835	24	98
Literary Prize	65418	0.034	0.182	0	1
Age at Death	65418	76.430	11.040	36	98

Notes: Bestseller lists are from the newspaper *La Stampa - Tuttolibri* (years 1975-2005) www.lastampa.it/archivio-storico. Information on dead authors are drawn from the *Biographical Dictionary of the Italians (Istituto dell'Enciclopedia Italiana – Treccani).*

[8] The Strega Prize is the most important Italian literary prize. It was launched in Rome in 1947 by writers Goffredo and Maria Bellonci (with the contribution of Guido Alberti, manufacturer of Strega liquor from which the prize took its name). The winner is chosen annually in July among books published between April 1 of the previous year and March 31 of the current year (see Ponzo and Scoppa, 2015, for further details on this prize). The Bagutta Prize was created in 1927 and is the oldest literary Italian award; it is assigned each year to the best piece of writing in any form: novels, short-story anthologies, poetry collections, memoirs. The Bancarella Prize, since 1952, goes to the best-selling book in the solar year, held in July in Pontremoli (Tuscany). The Campiello Prize was established in Venice in 1963 by the local Chamber of Commerce, to strengthen business relationships with the literary industry and to further promote Italian literature. The Viareggio Prize is a literary prize, first awarded in 1930 and named after the Tuscan city of Viareggio, that was conceived to rival the Milanese Bagutta Prize.

[9] The main publishers are: Mondadori, Rizzoli, Einaudi, Bompiani, Rusconi, Garzanti, Adelphi, Sellerio, Bur, Feltrinelli. The cumulated sales of the main publishers count for 78% of the total.

3. The Impact of Writers' Death on Book Sales: A Parametric Regression Discontinuity Approach

In this Section, to estimate the effect of an author's death on the sales of his/her books we adopt a Sharp Regression Discontinuity Design (RDD) (Imbens and Lemieux 2008; Angrist and Pischke 2009) and model the probability of an author of appearing in the bestseller list as a function of the time passed from the first appearance (our forcing or running variable in the jargon of RDD), using as control variables a number of individual characteristics. Then, we verify if the death of an author leads to a discontinuity in the function relating the bestseller list probability with time.

In general, a Regression Discontinuity Design allows to compare the outcomes of units just above the threshold with the outcomes of units just below the threshold. In our case, RDD compares our measures of book sales in the period just before and after the date of death (t_0) of an author. The time passed after the first appearance could in principle affect the sales of an author's book but under the assumption that the relationship between the outcome variable and time is continuous in a neighborhood of t_0, any jump in the dependent variable in proximity of the cutoff point can be interpreted as evidence of a treatment effect.

We first adopt a parametric approach estimating the following equation:

$$Bestseller_{it} = \beta_0 + \beta_1 AfterDeath_{it} + f(Time_{it}) + \boldsymbol{\beta_2} X_{it} + \lambda_t + u_{it} \qquad [1]$$

where the dependent variable $Bestseller_{it}$ is a dummy equal to one if the author *i* in week *t* is present in the bestseller list and equal to zero otherwise (alternatively, we use the *Points* earned in week *t* in the bestseller list), $AfterDeath_{it}$ is a dummy equal to one for a period of 6 months after the death of an author; $f(Time)$ is a flexible polynomial function of time – our forcing variable – centered at a writer's death date; X_{it} is a vector of author's characteristics (gender, age, age squared, literary prizes won, etc.) which could affect sales; λ_t is a set of monthly and yearly dummies; u_{it} is an error term. Therefore, in equation [1] the coefficient of interest is β_1 that represents the effect on the probability of entering in the bestseller list in the period immediately after the death of an author.

In the main analysis we estimate with a Linear Probability Model (while we use a Probit model in the Robustness Check Section). Since our main variable *After Death* is defined at the author's level, Standard Errors are allowed to cluster at the author's level and are robust to heteroskedasticity.

Estimation results are reported in Table 2. In column (1) we only use *After Death* and the forcing variable *Time* in linear form, without using any control variables. We find that after the death of an author the probability of appearing in the bestseller list increases of about 5.6 percentage points (p.p.), statistically significant at the 1 percent level (*t*-stat=3.39). With respect to the average probability of 5%, this corresponds to an increase of about 112%.

In column (2) we include as controls a number of potential determinants of the dependent variable: *Female*, *Age*, *Age Squared* and *Literary Prize*. The effect of female is not significant while the function between the dependent variable and age turns out to be concave, increasing until age 57.77 and decreasing thereafter. In addition, our estimates show a huge impact for Literary Prize (an increase of 22 p.p. after a Prize is won, which corresponds to an increase of 500% with respect to the average probability). More importantly, the impact of an author's death increases slightly to 5.9 p.p.

In column (3) we include a dummy for each month of the year and a dummy for each year from 1975 to 2005. Finally, in column (4) we add publisher fixed effects. Our results are confirmed: when we include all these controls, the impact on an author's death is about 5 p.p. (t-stats are around 3).

Table 2. The Probability of Entering in the Bestseller List After Death. Linear Probability Model

	(1)	(2)	(3)	(4)
After Death	0.056^{***}	0.059^{***}	0.049^{***}	0.052^{***}
	(0.017)	(0.017)	(0.017)	(0.017)
Time	-0.000^{**}	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Female		0.004	-0.002	-0.014
		(0.012)	(0.012)	(0.015)
Age		0.005^{**}	0.005^{**}	0.006^{**}
		(0.003)	(0.003)	(0.003)
Age Sq.		-0.000^{**}	-0.000^{**}	-0.000^{**}
		(0.000)	(0.000)	(0.000)
Literary Prize		0.222^{***}	0.204^{***}	0.198^{***}
		(0.034)	(0.034)	(0.035)
Month and Year Dummies	NO	NO	YES	YES
Publishers' Dummies	NO	NO	NO	YES
Observations	65418	65418	65418	65418
R^2	0.004	0.044	0.065	0.081

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

Different polynomial functions and interactions

The consequences of a misspecification of the functional form relating the outcome to the forcing variable are particularly serious in a RD design. We verify the robustness of our results to different polynomial trends, replicating the estimates in Table 2 first using a quadratic function of *Time* (columns 1-4 of Table 3) and then using a cubic function of *Time* (columns 5-8 of Table 3).[10] In both cases we show that the effect of an author's death remains strong, almost of the same magnitude and highly statistically significant (t-stats around 3): after the death of an author the probability of entering in the bestseller list increases of about 5-6 p.p.

[10] Gelman and Imbens (2019) show that in a RD design it is not appropriate to use higher order polynomials.

Table 3. Using a Quadratic and Cubic Function of Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After Death	0.052***	0.051***	0.048***	0.046***	0.063***	0.061***	0.054***	0.054***
	(0.017)	(0.017)	(0.016)	(0.016)	(0.019)	(0.018)	(0.017)	(0.017)
Time	YES	YES	YES	YES	YES	YES	YES	YES
Time Sq.	YES	YES	YES	YES	YES	YES	YES	YES
Time Cub.					YES	YES	YES	YES
Observations	65418	65418	65418	65418	65418	65418	65418	65418
R^2	0.004	0.044	0.065	0.081	0.004	0.045	0.066	0.081

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Controls in columns (1)-(4) and in columns (5)-(8) are the same of Table 2. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In Table 4 we re-estimate our model with basic controls (odd columns) and with the full range of controls (even columns) and interact our treatment variable *After Death*, with, respectively, the linear, quadratic and cubic function of time to verify if a different pattern of the forcing variable around the cutoff might be erroneously exchanged for a discontinuity. Results of Table 4 confirm that *After Death* has a strong impact on sales of books even when we control for a polynomial of time and for all the interaction terms.

Table 4. Robustness to Different Functional Forms across the Cutoff

	(1)	(2)	(3)	(4)	(5)	(6)
After Death	0.066***	0.055***	0.042***	0.033**	0.041***	0.028*
	(0.018)	(0.019)	(0.015)	(0.016)	(0.014)	(0.017)
Time	-0.000	0.000	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Time*After Death	-0.001	-0.001	0.007**	0.009**	0.019***	0.021***
	(0.001)	(0.001)	(0.003)	(0.004)	(0.007)	(0.007)
Time Sq.			0.000	0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)	(0.000)
Time Sq.*After Death			-0.000***	-0.000***	-0.002**	-0.002**
			(0.000)	(0.000)	(0.001)	(0.001)
Time Cub.					-0.000	-0.000
					(0.000)	(0.000)
Time Cub.*After Death					0.000**	0.000**
					(0.000)	(0.000)
Controls	Basic	All	Basic	All	Basic	All
Observations	65418	65418	65418	65418	65418	65418
R^2	0.044	0.065	0.044	0.066	0.045	0.066

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Basic controls include: Female, Age, Age Squared, Literary Prize. In even columns we add month and year dummies. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

Local Windows

Notwithstanding our checks, in principle our estimates could still be biased by an incorrect specification of the functional form relating the outcome variable to time. To avoid this risk, it is usual to estimate the model using only data in a neighborhood around the discontinuity (Imbens and Lemieux, 2008). In practice, in our analysis we need to compare the sales of books in some time window around the date of death of writers.

In columns (1) and (2) of Table 5 (with basic and with full controls, respectively) for each writer we consider only data for a period of 150 weeks – approximately 3 years – before and after his/her death (28,000 obs.). For this sample, the impact we estimate for the probability of entering in the bestseller list after the writer's death corresponds to an increase of about 3.5 p.p., statistically significant at the 1 percent level. In columns (3) and (4) we only focus on data of 100 weeks (2 years) before and after the cutoff. We find an impact of about 5 p.p. In columns (5) and (6) we focus on the data relative to the 50 weeks (1 year) before and after the cutoff. We find an impact of about 6.5 p.p. Finally, in columns (7) and (8) we take into account only data for 13 weeks (3 months) before and after the threshold finding an impact of about 8 p.p.

All the estimated effects focusing on symmetric windows around the threshold are highly statistically significant, notwithstanding the considerable lower number of observations. The magnitude of the impact turns out to be stronger the shorter the time period we consider, suggesting an immediate response of consumers after the event and a decline thereafter.

Table 5. RD Estimates Using Different Local Windows

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Window	±150 weeks	±150 weeks	±100 weeks	±100 weeks	±50 weeks	±50 weeks	±13 weeks	±13 weeks
After Death	0.036***	0.034***	0.055***	0.052***	0.065***	0.065***	0.077***	0.081***
	(0.012)	(0.011)	(0.015)	(0.014)	(0.022)	(0.022)	(0.026)	(0.029)
Controls	Basic	All	Basic	All	Basic	All	Basic	All
Observations	28609	28609	20918	20918	10700	10700	2931	2931
R^2	0.002	0.221	0.005	0.258	0.007	0.299	0.017	0.561

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Basic controls include: Time, Female, Age, Age Squared, Time, Literary Prize. All controls include in addition to the former also month and year dummies. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

An Alternative Outcome: Points in the Bestseller List

In the previous analyses we have considered as outcome variable the probability of being in the bestseller list. Another outcome we have available, with more detailed information, is the number of points in the bestseller list reflecting the amount of copies sold each week, where 100 points are awarded to the first ranked book in each week and the other points are assigned to books proportionally to their volume of sales.

In Table 6 we replicate specifications in Tables 2 and 3 (for the latter, only columns 1-4) and we find that after death the bestseller points earned by an author increase of about 3.1-3.5 in each week, which corresponds to about 0.30 Standard Deviations of the dependent variable. In Table 7 we replicate estimates in Table 5 using *Bestseller Points* as a dependent variable and again we find a strong impact, ranging from 2.1 points, when we consider larger windows, to 6.4 points, when we consider narrower windows.

Table 6. The Impact of a Writer's Death on Bestseller Points

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
After Death	3.380***	3.489***	3.229**	3.349**	3.396***	3.349***	3.198**	3.101**
	(1.227)	(1.262)	(1.266)	(1.280)	(1.250)	(1.256)	(1.226)	(1.211)
Observations	65418	65418	65418	65418	65418	65418	65418	65418
R^2	0.004	0.031	0.045	0.062	0.004	0.031	0.045	0.062

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller Points*. The specifications in columns (1)-(4) are the same of Table 2; the specifications in columns (5)-(8) are the same of Table 3 (columns 1-4). Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

Table 7. Bestseller Points and Local Windows

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Window	±150 weeks	±150 weeks	±100 weeks	±100 weeks	±50 weeks	±50 weeks	±13 weeks	±13 weeks
After Death	2.141***	2.012***	3.145***	3.003***	4.469**	4.455**	6.291***	6.483***
	(0.765)	(0.728)	(1.089)	(1.079)	(1.854)	(1.883)	(1.968)	(2.287)
Controls	Basic	All	Basic	All	Basic	All	Basic	All
Observations	28609	28609	20918	20918	10700	10700	2931	2931
R^2	0.003	0.167	0.005	0.210	0.008	0.256	0.014	0.492

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller Points*. Basic controls include: Time, Female, Age, Age Squared, Time, Literary Prize. All controls include in addition to the former also month and year dummies. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

4. A Non-Parametric Regression Discontinuity Approach

In this Section, as further robustness checks, we follow a Non-Parametric RD approach. This approach tries to produce reliable estimates focusing on small samples to the right and to the left of the cutoff point ("*bandwidths*"), through the so-called Local Linear Regressions and Polynomial Regressions that represent a sort of Weighted Least Squares estimations of equation [1], where the weights are larger the closer the observations are to the cutoff (using different functions, or *kernels*, to determine the weights).

We use the procedures implemented by Calonico, Cattaneo, Farrell, and Titiunik (2017; 2019) and Calonico, Cattaneo, and Farrell (2018)[11] and experiment with different bandwidths, different kernels and different polynomial orders.

Our non-parametric estimates, using *Bestseller List* as the outcome variable, are reported in Table 8. In column (1) we use a uniform kernel, a linear polynomial and the Mean Squared Error optimal bandwidth selection ("Mserd") – a data-driven bandwidth selection that optimizes the Mean Squared Error – that selects a bandwidth of 22 weeks before and 22 after the cutoff: the estimated effect is 8.3 p.p., slightly higher in magnitude but qualitatively similar to the parametric estimates of Table 5.

[11] We refer to these works for further details on the methodology.

The estimated effect turns out to be very similar (ranging from 5 to 9 p.p.) when we use different bandwidths (Msetwo,[12] or manually setting windows to ±50 and ±100 from the cutoff) in columns 2-4; or using different kernels – respectively, Triangular and Epanechnikov – in columns 5 and 6; or using a second and third polynomial order (columns 7 and 8).

We have also carried out the same estimates including our basic covariates that we used in the parametric analysis. Again, we find quite similar results (estimates not reported to save space).

Table 8. RD Non-Parametric Estimates of the Impact of a Writer's Death. Dependent Variable: Bestseller List

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RD Estimate	0.083***	0.090***	0.063***	0.053***	0.083***	0.084***	0.090***	0.096***
	(0.015)	(0.013)	(0.010)	(0.006)	(0.015)	(0.015)	(0.017)	(0.019)
Observations	76090	76090	76090	76090	76090	76090	76090	76090
Bandwidth Type	Mserd	Msetwo	±50	±100	Mserd	Mserd	Mserd	Mserd
Kernel	Uniform	Uniform	Uniform	Uniform	Triangul.	Epanech.	Uniform	Uniform
Eff. obs – Left of c	2355	11219	5265	10325	3199	2884	4338	6392
Eff. obs – Right of c	2495	2285	5435	10593	3335	3020	4490	6585
Order Polynomial	1	1	1	1	1	1	2	3

Notes: The Table reports RD estimates using local polynomial regressions with robust bias-corrected inference procedures. The estimates are implemented using the Stata program *rdrobust* by Calonico, Cattaneo, Farrell, and Titiunik. The dependent variable is *Bestseller List*. *Mserd* indicates one common MSE-optimal bandwidth selector for the RD treatment effect estimator. *Msetwo* indicates Two different MSE-optimal bandwidth selectors (below and above the cutoff).

In Table 9 we run the same estimates using *Bestseller Points* as a dependent variable and again we find very similar and consistent results: the number of bestseller points earned in the period following the death of a writer is estimated in the range 3-7, according to the choices of bandwidths and kernels.

Table 9. RD Non-Parametric Estimates of the Impact of a Writer's Death. Dependent Variable: Bestseller Points

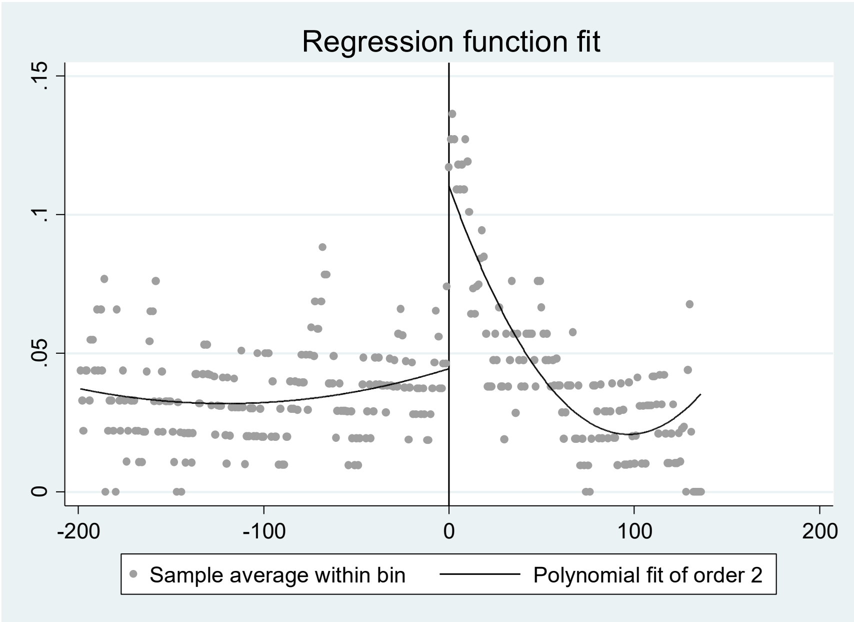
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RD Estimate	5.858***	7.537***	4.275***	3.068***	5.926***	5.944***	6.689***	6.988***
	(1.300)	(1.190)	(0.682)	(0.411)	(1.291)	(1.275)	(1.310)	(1.499)
Observations	76090	76090	76090	76090	76090	76090	76090	76090
Bandwidth Type	Mserd	Msetwo	±50	±100	Mserd	Mserd	Mserd	Mserd
Kernel	Uniform	Uniform	Uniform	Uniform	Triangular	Epanechnikov	Uniform	Uniform
Eff. obs – Left of c	2036	27545	5265	10325	2884	2567	4441	6290
Eff. obs – Right of c	2180	1968	5435	10593	3020	2705	4595	6481
Order Polynomial	1	1	1	1	1	1	2	3

Notes: The Table reports RD estimates using local polynomial regression with Robust Bias-Corrected Confidence Intervals and Inference Procedures. The estimates are implemented using the Stata program *rdrobust* by Calonico, Cattaneo, Farrell, and Titiunik. The dependent variable is *Bestseller Points*. Mserd indicates one common MSE-optimal bandwidth selector for the RD treatment effect estimator. *Msetwo* indicates Two different MSE-optimal bandwidth selectors (below and above the cutoff).

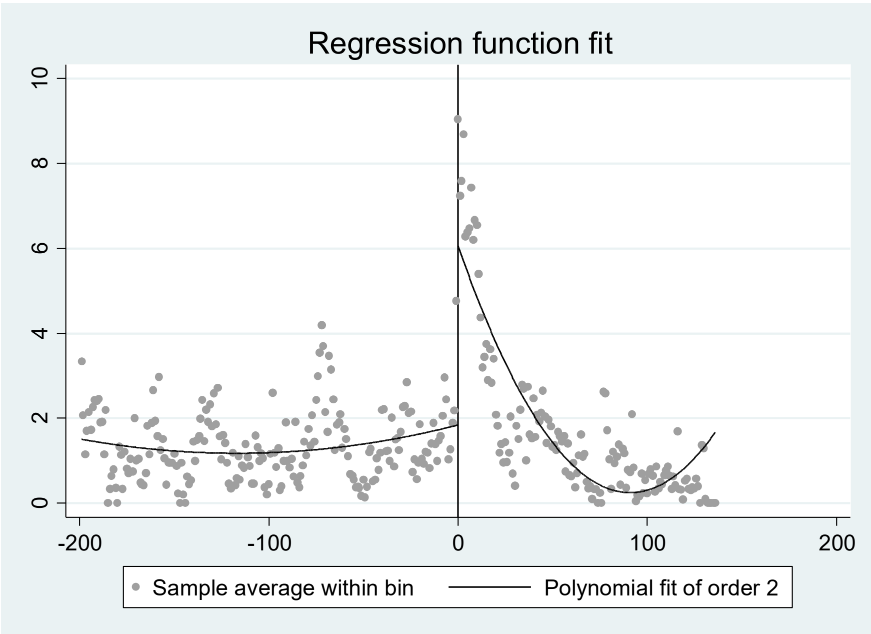
One of the main advantage of the Regression Discontinuity Design is that it allows a transparent graphical analysis. In Figure 1 we represent with dots the sample average within each bin of Bestseller List (Panel A) or Bestseller Points (Panel B) against time in weeks. We also represent with a continuous line the

[12] Two different MSE-optimal bandwidth selectors (below and above the cutoff).

predicted values from a second-order polynomial, estimated separately on each side of the cutoff point. The vertical line at time 0 denotes the week of death of each writer. In both panels of Figure 1 it clearly emerges a marked jump in the relationship between the outcomes and the time variable in the proximity of the threshold.



Panel (A)



Panel (B)

Figure 1. The Probability of Entering in the Bestseller List (Panel A) and Bestseller Points (Panel B) as a Function of Time

The vertical line at Time 0 denotes the death of an author

4.1. RDD Validity Checks

In this Section we run two validity tests to investigate if the identifying assumptions of the RD approach are satisfied. Two key RDD assumptions are that: a) there is no "manipulation" of the forcing variable and so the cutoff provides exogenous variations in the treatment; b) observable and unobservable characteristics do not vary discontinuously at the cutoff.

Notwithstanding in our context manipulation of the forcing variable is completely unrealistic, to test for the existence of manipulation we plot a histogram of the density of the forcing variable around the zero cutoff as suggested by McCrary (2008). The histogram does not show any evidence of jumps. In addition, we run the McCrary test, a formal RD manipulation test using local polynomial density estimation and find a coefficient of 0.622, which, with a *p*-value=0.53, is far from being significant.

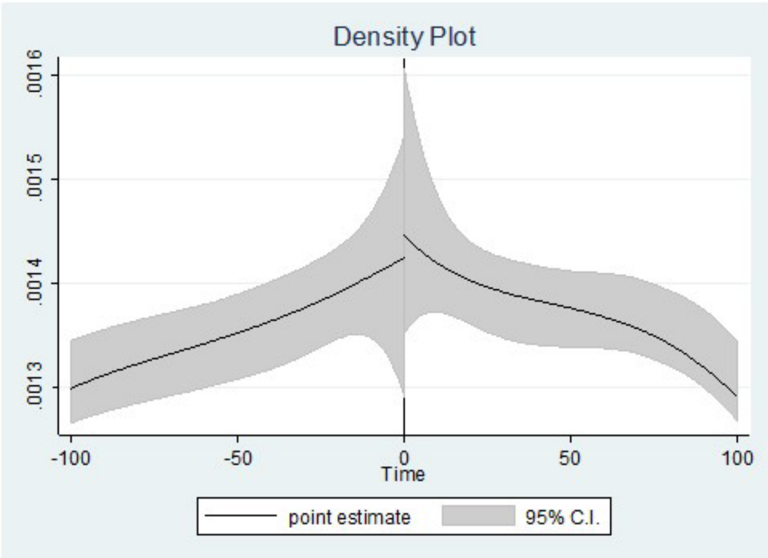


Figure 2. The Density of the Forcing Variable

As regards the second validity check, we focus on pre-determined variables (female, age, literary prizes, large publishers) and we test the continuity of these variables at the threshold to control whether some kind of discontinuity emerges in proximity of an author's death. In practice, we estimate an equation analogous to equation (1), where the dependent variable is, in turn, a predetermined characteristic.

Table 10 reports the coefficients on *After Death* estimated using local polynomial regressions.[13] Overall, our estimates show that the death of an author is not associated with any discontinuity in our predetermined characteristics, confirming that the effect estimated in our main analysis is not due to any spurious correlation between the death of an author and other factors.

[13] We obtain very similar results if we run parametric regressions.

Table 10. No Discontinuity in the Control Variables. RD estimates using Local Polynomial Regressions

	(1) Female	(2) Age	(3) Literary Prize	(4) Mondadori	(5) Rizzoli	(6) Einaudi	(7) Bompiani
RD Estimate	-0.0045	0.0227	0.0004	-0.0045	-0.0038	0.0134	-0.0036
	(0.0195)	(0.565)	(0.0074)	(0.0151)	(0.0143)	(0.0137)	(0.0121)
Observations	76090	76090	76090	76090	76090	76090	76090
Bandwidth Type	Mserd	Mserd	Mserd	Mserd	Mserd	Mserd	Mserd
Kernel	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform
Eff. obs – Left of c	3094	3511	3926	5162	5059	4956	4441
Eff. obs – Right of c	3230	3650	4070	5330	5225	5120	4595
Order Polynomial	1	1	1	1	1	1	1

Notes: The Table reports estimates using local polynomial regressions with Robust Bias-Corrected Confidence Intervals and Inference Procedures. The estimates are implemented using the Stata program *rdrobust* by Calonico, Cattaneo, Farrell, and Titiunik. The dependent variable is reported at the top of each column. Mserd indicates one common MSE-optimal bandwidth selector for the RD treatment effect estimator.

5. Robustness Checks

In this Section we run a number of robustness checks to verify if our results are driven by some specific choices we have made in the analysis. To check the robustness of our results we will use different estimation methods, different outcome variables and different samples.

Estimates with Author Fixed Effects

In Table 11 we estimate our model controlling for author fixed effects using the parametric approach. In this analysis we identify the effect of our interest not comparing different authors but rather comparing for each author the sales of his/her books in the period before and after his/her death. We estimate the same specifications as in columns (1)-(4) of Table 2, excluding only time invariant variables such as gender. The impact remains quite strong and very stable, around 6 p.p.

Table 11. The Probability of Entering in the Bestseller List After Death with Author Fixed Effects. Linear Probability Model

	(1)	(2)	(3)
After Death	0.061^{***}	0.063^{***}	0.060^{***}
	(0.017)	(0.017)	(0.017)
Controls	No	Basic	All
Author Fixed Effects	YES	YES	YES
Observations	65418	65418	65418
R^2	0.128	0.128	0.157

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Basic controls include: Time, Age, Age Squared, Literary Prizes. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In our analyses so far we have estimated with a Linear Probability Model. In Table 12 instead we estimate with a Probit Estimator and calculate the corresponding Average Marginal Effects. Again, our results are confirmed: after the death of an author the probability of entering in the bestseller list increases of about 4-5 p.p.

Table 12. Probit Estimates for the Probability of Entering in the Bestseller List. Average Marginal Effects

	(1)	(2)	(3)	(4)
After Death	0.048***	0.051***	0.044***	0.047***
	(0.012)	(0.011)	(0.011)	(0.011)
Controls	None	Basic	All	All
Observations	65418	65418	65418	65418
Pseudo R^2	0.002	0.030	0.039	0.081

Notes: The Table reports Average Marginal Effects of Probit estimates. The dependent variable is *Bestseller*. Basic controls include: Time, Female, Age, Age Squared, Literary Prizes. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

We then consider two additional outcome variables. In Panel (A) of Table 13 we use as an outcome variable the number of books entered in the bestseller list. We find that the number of books increases of about 0.09 in the weeks immediately after the death of a writer (statistically significant at the 1 percent level).

In Panel (B) of Table 13 we examine the impact on the ranking in the bestseller list, using a variable based on the position in the list, *Ranking*, defined in the following way: 4=positions 1-5; 3=positions 6-10; 2=positions 11-15; 1=positions 16-20; 0=Not in the bestseller list. We find that the ranking position significantly improves in the period after death.

Table 13. The Impact on the Number of Books in the Bestseller List and Ranking Positions

Panel (A): Number of Books in the Bestseller List

	(1)	(2)	(3)	(4)	(5)
After Death	0.088^{***}	0.091^{***}	0.081^{***}	0.084^{***}	0.090^{***}
	(0.027)	(0.027)	(0.028)	(0.028)	(0.029)
Controls	Only Time	Basic	All	All	All
Publisher FE	No	No	No	Yes	No
Author FE	No	No	No	No	Yes
Observations	65418	65418	65418	65418	65418
R^2	0.006	0.041	0.063	0.079	0.162

Panel (B): Ranking Position

	(1)	(2)	(3)	(4)	(5)
After Death	0.182^{***}	0.190^{***}	0.167^{***}	0.176^{***}	0.202^{***}
	(0.057)	(0.057)	(0.058)	(0.058)	(0.060)
Controls	Only Time	Basic	All	All	All
Publisher FE	No	No	No	Yes	No
Author FE	No	No	No	No	Yes
Observations	65418	65418	65418	65418	65418
R^2	0.003	0.042	0.058	0.075	0.155

Notes: The Table reports OLS estimates. The dependent variable in Panel (A) is the *Number of Books in the Bestseller List*. The dependent variable in Panel (B) is the *Ranking* (4=positions 1-5; 3=positions 6-10; 2=positions 11-15; 1=positions 16-20; 0 not in the bestseller list). Basic controls include: Time, Female, Age, Age Squared, Literary Prizes. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In our main analysis we have considered only the authors died in the period 1976-2005. We now analyze if changing the sample of authors can modify our results. In Table 14 we consider all the authors entered in the bestseller lists and who have died in the period 1976-2021 (223 authors). Notice that in this alternative sample our treatment variable does not change (equal to one for the three months after the death of authors) but only increases the number of authors in the control group. Our estimates are confirmed: the coefficients turn out to be slightly higher in magnitude (about 5-6 p.p.) and statistical significance.[14]

Table 14. Estimates on a Different Sample of Authors (Died in the Period 1976-2021)

	(1)	(2)	(3)	(4)	(5)
After Death	0.057^{***}	0.060^{***}	0.055^{***}	0.056^{***}	0.060^{***}
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Controls	Only Time	Basic	All	All	All
Publisher FE	No	No	No	Yes	No
Author FE	No	No	No	No	Yes
Observations	68218	68218	68218	68218	68218
R^2	0.004	0.045	0.065	0.079	0.169

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Basic controls include: Time, Female, Age, Age Squared, Literary Prizes. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

[14] Furthermore, we obtain similar results if in the control group we consider also the authors still alive (estimates not reported).

6. Heterogeneity and Possible Mechanisms at Work

In this Section, to shed some light on the mechanisms at work we investigate if the impact of the death of an author is different according to age at death, to the reputation of the authors or to the prominence of the news.

First of all, we consider the age at death of writers and split the sample in two subsamples: authors died at age 65 or below and authors died at age higher than 65. The idea is that emotions of readers might be more intense in case of premature deaths of writers. In columns 1 and 2 of Table 15 we estimate our preferred specification (column 3, Table 2, without and with author fixed effects) on the sample of authors died prematurely (only 22 authors in this sample): we find that the impact of our interest is much larger, about 10 p.p. In columns 3 and 4 of Table 15 we estimate the same specifications on the sample of authors who died after age 65. In this case the impact is about the half, around 4.5 p.p. We find very similar results if we estimate on the whole sample and include an interaction term between *Premature Death* and *After Death*. Results are also similar if we split the sample at age 70.

These findings suggest that consumers are likely to react emotionally with respect to premature deaths of authors, while consumers turn out to be less reactive when authors die older.

Table 15. Heterogeneous Effects on the Basis of Age at Death. RD Estimates

Age at Death	Less than 65			More than 65
	(1)	(2)	(3)	(4)
After Death	0.102*	0.107**	0.044***	0.032**
	(0.052)	(0.051)	(0.015)	(0.016)
Controls	Basic	All	Basic	All
Observations	9608	9608	55810	55810
R^2	0.052	0.088	0.045	0.070

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Basic controls include: Time, Female, Age, Age Squared, Literary Prize. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

Next we consider the authors' fame. To this aim, using our data on bestseller lists we calculate for each author the number of bestseller points earned before his/her death, *Fame Index*. We then split the sample according to the median value of *Fame Index* (279.8).

In columns (1) and (2) of Table 16 we estimate on the sample of less famous authors: we find a very small effect, about 1%, not statistically significant at conventional levels (p-value=0.16). In columns (3) and (4) of Table 16 instead we estimate on the sample of famous authors: we find a larger impact, around 10-11 percentage points. This result points to media attention as a possible driver of the uncovered effect: for famous authors, for whom news are prominent, media attention is particularly high and this affects the demand of consumers.

Table 16. Heterogeneous Effects on Author's Fame. RD Estimates

	Less Famous Authors		Famous Authors	
	(1)	(2)	(3)	(4)
After Death	0.016	0.011	0.108***	0.101***
	(0.010)	(0.008)	(0.033)	(0.033)
Controls	Basic	All	Basic	All
Observations	32692	32692	32726	32726
R^2	0.004	0.023	0.049	0.073

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Basic controls include: Time, Female, Age, Age Squared, Literary Prize. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

To better understand this issue, we have exploited the historical archive of *La Stampa* and gathered data on the following aspects: 1) whether the news of the death of an author appeared on the front page or not in the following week; 2) whether the news appeared on the first page of the Cultural Section; 3) the total number of articles related to the event.

On the basis of these variables we estimate a model (using a specification with all the controls, column 3 of Table 2) for the sample of authors whose death news appeared on the front page (column 1, Table 17) and for the remaining authors (column 2, Table 17). We find a very large effect for authors covered on the front page (+13.3 p.p.) whereas we find no effect for the other authors.

We obtain similar results when we distinguish authors on the basis of the appearance of the event in the first cultural page (columns 3 and 4, Table 17). Finally, in columns (5) and (6) we split our sample on the basis of the total number of articles (wherever in the newspaper) related to the death event, whether the number of articles is above or below 3. Again, we find a huge effect (+12.2 p.p.) when the death of an author has been covered more extensively in the media, while no effect is found when the event has received little or no attention.

All these findings are confirmed when we estimate on the whole sample and use an interaction term between *After Death* and an indicator for the prominence of the news coverage (estimates not reported).

Therefore, the evidence from Tables 16 and 17 shows that media attention (and the related author's fame) is crucial in determining the impact of a writer's death on consumers' decisions.

Table 17. Relevance of the News. RD Estimates

	Front Page	No Front Page	First Cultural Page	No First Cultural Page	#Newspaper articles<=3	#Newspaper articles>3
	(1)	(2)	(3)	(4)	(5)	(6)
After Death	0.133***	0.012	0.097***	-0.011	0.001	0.122***
	(0.044)	(0.011)	(0.028)	(0.008)	(0.011)	(0.035)
Controls	All	All	All	All	All	All
Observations	19886	45532	28989	36429	40818	24600
R^2	0.078	0.059	0.062	0.079	0.060	0.080

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

Finally, we also try to verify if the impact is affected by the relevance of the publishers, that could be related to advertising campaign triggered by the author's death. To this aim, we distinguish between small and large publishers on the basis of the weeks of appearance of each publisher in the bestseller list, splitting the sample at the median value. We analyze the sample with small publishers in columns (1) and (2) of Table 18 and the sample with large publishers in columns (3) and (4). We find that the effect of *After Death* is larger (around 7-8 p.p.) when we consider small publishers, while it ranges between 4-5 p.p. for large publishers. The difference turns out to be not statistically significant when we estimate on the whole sample and include an interaction between *After Death* and *Large Publisher*. Therefore, the resources of a publisher do not seem to affect significantly the consumers' reactions to the event.

Table 18. Large and Small Publishers. RD Estimates

	Small Publishers			Large Publishers
	(1)	(2)	(3)	(4)
After Death	0.078^{**}	0.070^{**}	0.047^{***}	0.039^{*}
	(0.033)	(0.032)	(0.018)	(0.020)
Controls	Basic	All	Basic	All
Observations	21184	21184	44234	44234
R^2	0.036	0.060	0.049	0.048

Notes: The Table reports OLS estimates. The dependent variable is *Bestseller*. Basic controls include: Time, Female, Age, Age Squared, Literary Prize. Standard errors (reported in parentheses) are corrected for heteroskedasticity and for clustering at the author level. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

7. Concluding Remarks

In the standard economic model spurious information as the death of an author should have no impact on the sales of his/her books, since consumers are assumed to be fully rational, emotionless and always using all the available information in their economic decisions. In contrast, if – as suggested by behavioral economics – individuals react to emotions and exploit information that are made salient by an event such as the death of an author while neglecting them in other periods because of limited attention, then the decision to buy a book may be affected by the announcement of an author's death. This is in line with several studies showing that emotions might affect individual decisions in many contexts.

In this paper using data on bestseller lists at week level for 30 years (1975-2005) we have carried out an empirical analysis to study the impact of an author's death on the sales of his/her books. We have used a Regression Discontinuity Design and showed that in the period immediately following his/her death the probability of entering in the bestseller list increases of about 4-5 percentage points, which corresponds to an impact of more than 100%. A number of robustness checks – controlling for several factors potentially affecting sales, using parametric and non-parametric estimations, changing the time window around the date of death, the estimation method, the sample and using different outcome variables – has confirmed a strong impact of an author's death on book sales.

In the attempt to investigate possible mechanisms driving the uncovered effect, we have found a much larger impact for the authors dying at an early age – which suggests that emotions may play an important role in the reactions of consumers. A larger effect is also associated to more famous writers and to situations in which the news of the death event is covered by media more extensively, suggesting that media attention, raising the awareness of readers to the life and works of an author, represents also a crucial factor in stimulating consumers' demand.

This analysis has been conducted for books thanks to the availability of data, but we think our findings can be extended to other fields. For example, the markets for cultural products such as films, musical products, artistic works are likely strongly affected when an author dies. Similarly, when a firm, a manager, an entrepreneur, etc. for some reason end up on the front pages of newspapers or in social networks, the media will capture the attention of consumers and the markets of the related products will likely be affected.

Future research could investigate this kind of effects on some other fields and would benefit from using more accurate data on effective sales rather than on the presence in the bestseller lists or related rankings.

References

Angrist, J., & Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3), 442-451.

Calonico, S., M. D. Cattaneo, & M. H. Farrell. 2018. On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference. *Journal of the American Statistical Association*, 113(522): 767-779.

Calonico, S., M. D. Cattaneo, M. H. Farrell, & R. Titiunik. 2017. rdrobust: Software for Regression Discontinuity Designs. *Stata Journal*, 17(2): 372-404.

Card, D., & Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *Quarterly Journal of Economics*, 126(1), 103-143.

Coase, R., (1972). Durability and Monopoly, *Journal of Law and Economics* 15 (April): 143–149.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, 47(2), 315-72.

Ekelund, R. B., Ressler, R. W., & Watson, J. K. (2000). The "Death-Effect" in Art Prices: A Demand-Side Exploration. *Journal of Cultural Economics*, 24(4), 283-300.

Enciclopedia Istituto Treccani, *Dizionario Biografico degli Italiani*, 2020.

Frey, B. S., & Gullo, A. (2020). Sic transit gloria mundi: What remains of famous economists after their deaths? *Scientometrics*, 123.

Gabaix, X., 2019. Behavioral inattention. In: Bernheim, D., DellaVigna, S., Laibson, D. (Eds.), *Handbook of Behavioral Economics*, vol. 2. North-Holland, Amsterdam, Pages 261-343.

Gelman, A., & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, *37*(3), 447-456.

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

Hsee, C. K., & Rottenstreich, Y. (2004). Music, pandas, and muggers: On the affective psychology of value. *Journal of Experimental Psychology: General*, 133, 23–30.

Huberman, G., & Regev, T., 2001. Contagious Speculation and a Cure for Cancer: A Nonevent That Made Stock Prices Soar. *Journal of Finance*, 56(1): 387–96.

Imbens, G., & Lemieux, T., (2008), Regression Discontinuity Designs: A Guide to Practice, *Journal of Econometrics*, 142 (2), 615–635.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *American Economic Review*, 90(2), 426-432.

Loewenstein, G., & J. Lerner. (2003). The role of affect in decision making, in Davidson, R. J., Scherer, K. R., & Goldsmith, H. H. (2003). *Handbook of Affective Sciences*, 619-642.

Maddison, D., & J. Pedersen, A. (2008). The death effect in art prices: evidence from Denmark. *Applied Economics*, *40*(14), 1789-1793.

Matheson, V. A., & Baade, R. A. (2004). 'Death effect' on collectible prices. *Applied Economics*, *36*(11), 1151-1155.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2), 698-714.

Nelson, P. (1970). Information and consumer behavior, *Journal of Political Economy*, 78, pp. 311-329.

Ponzo, M., & Scoppa, V. (2015). Experts' awards and economic success: evidence from an Italian literary prize. *Journal of Cultural Economics*, *39*(4), 341-367.

Radford, S. K., & Bloch, P. H. (2013). Consumers' online responses to the death of a celebrity. *Marketing Letters*, 24(1), 43-55.

Rick, S., & Loewenstein, G. (2008). The role of emotion in economic behavior. *Handbook of emotions*, 3, 138-158.

Simon, H. A. (1986). The role of attention in cognition (pp. 105-115) in *The Brain, Cognition, and Education* (eds: Sarah L. Friedman, Kenneth A. Klivington, Rita W. Peterson), New York: Academic Press.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*. New York: Cambridge University Press, pp. 397-98.

Stanton, S. J., Reeck, C., Huettel, S. A., & LaBar, K. S. (2014). Effects of induced moods on economic choices. *Judgment and Decision Making*, 9(2), 167.

Thaler, R. H. (2016). Behavioral economics: Past, present, and future. *American Economic Review*, 106(7), 1577-1600.

Ursprung, H. W., & Wiermann, C. (2011). Reputation, price, and death: An empirical analysis of art price formation. *Economic Inquiry*, *49*(3), 697-715.