Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis BSc2 Econometrics / Economics

Immediate Market Response to Earnings Surprises: A Regression Discontinuity Analysis

Xander van der Meer (579204)

Supervisor:	Jens Klooster
Second assessor:	Dr. Eoghan O'Neill.
Date final version:	1 July 2024

Abstract

Earning announcements are known to move stock prices. This paper analyzes the effect of earnings surprises on the excess stock returns of the first trading day after the announcement. Utilizing a regression discontinuity design, we show a positive and significant treatment effect, with positive earnings surprises on the stock return. Also, earnings surprises and stock returns are positively correlated. The results are robust to the inclusion of additional covariates.

Contents

1 Introduction

Earnings publications can lead to share price movements. For instance, following Dinsey's earnings announcement, a headline read: "Disney Stock Limps To Worst Day Since 2022 Despite Surprise Disney+ Profit" (Saul, 2017). Similarly, Tesla stock dropped 3% following their earnings announcement, which was below analysts' expectations (Ruffino, 2024). The reverse can also be true. For example, Spotify's share price increased by 5% following unexpectedly high earnings (Bary, 2024). Public companies disclose their reported earnings every quarter, and analysts make forecasts about these earnings. When a company's reported earnings deviate from these forecasts, this is called an earnings surprise. The examples provided illustrate that earnings surprises immediately affect share prices. Despite extensive research by Barber et al. (2013), Frazzini and Lamont (2007), and Skinner and Sloan (2002) on the longer-term effects of earnings surprises, hardly any research exists on the immediate impact. This paper fills the gap in the literature by analyzing this immediate effect.

A benefit of only looking at the trading day following the publication is to isolate the effect of the earnings surprise. When looking at a longer time horizon, other factors will likely influence the share price. Next to the longer time horizon, research that has been done only uses simple linear regression instead of the more advanced econometric methods that this paper uses.

Research on the immediate effect is relevant in several ways. Firstly, it contributes to the understanding of market efficiency. Market efficiency is a central topic within the finance field. In an efficient market, all known information is incorporated into stock prices (Ţiţan, 2015). There are three essential assumptions required for this to hold. Firstly, investors are assumed to be rational. Secondly, if investors are not rational, their trades are random. Finally, influence from irrational investors is eliminated by arbitrageurs (Naseer and bin Tariq, 2016). There are three forms of market efficiency: weak, semi-strong and strong. Under the weak form, only historical prices are incorporated into the share price. With semi-strong market efficiency, the share price reflects all public information. Finally, strong market efficiency suggests that all public and private information is reflected by the share price (Naseer and bin Tariq, 2016). The different forms of market efficiency imply different reactions to earnings surprises. Under strong market efficiency, one would expect no change in the share price as the information is already reflected in the share price. With semi-strong market efficiency, the share price would quickly adjust to incorporate the new information. Finally, under the weak form, the market may need some time to adjust.

Secondly, the results give insight into investor behaviour, which is relevant to the behavioural finance field. Skiba and Skiba (2017) found that individual investors and even institutional investors are irrational. This violates the assumptions of the efficient market hypothesis. Different behavioural biases influence investors' choices. Prospect theory is a theory that applies to investors. According to prospect theory, people are more sensitive to the pain of losing than the pleasure of gaining an equivalent amount (Kahneman and Tversky, 1979). People evaluate outcomes relative to a reference point. For earnings announcements, this reference point could be the analysts' forecasts. This theory suggests that people react more heavily to negative earnings surprises than positive ones. This can cause a discontinuity at the threshold between the two. This conflicts with conventional economic theory that assumes that people are perfectly rational (Frank and Cartwright, 2013). If this were true, there is expected to be a continuous relationship between earnings surprises and stock returns.

Finally, portfolio managers can use the results to improve risk management. Earning announcements bring additional volatility to share prices Patell and Wolfson (1979). This can present an additional risk for portfolio managers. Knowing how share prices will react is valuable information, as it could be that additional hedging is necessary when earnings are announced to mitigate the risk of earnings surprises.

This leads to the following research question: How is the excess return of a stock over the S&P 500 return affected by an earning surprise on the first trading day following the earnings publication? The reason for looking at the excess returns over the S&P 500 returns is to isolate the effect of the earnings surprise from the general market movement. There are three subquestions to answer this question. Firstly, what is the effect on excess stock returns when earnings exceed expectations? This question looks into the effect on stock returns for the binary event of exceeding expectations; this gives a general understanding of the overall effect without looking at the magnitude of the earnings surprise. Secondly, how does the size of the earnings surprise affect the excess stock returns? The size of the earnings surprise is likely related to the size of the excess returns, but this could be in different ways. Thirdly, is the effect of earnings surprises on stock returns dependent on firm-specific characteristics? Firm factors such as size and industry can affect the reaction to earnings surprises. In specific sectors, forecast errors by analysts are more prominent than in other sectors Brown (1997), and this can change investor behaviour.

Sharp Regression discontinuity (RD) design is the method that is used in this paper. This method evaluates the treatment effect corresponding to a binary treatment received after the forcing variable crosses a threshold (Imbens and Lemieux, 2008). In this case, the treatment is outperforming the expectations of the analysts. The threshold is where companies report the same earnings as the forecast. A surprise ratio of the actual earnings divided by the expected earnings is constructed and serves as the forcing variable. The method fits local polynomial regressions on both sides of the threshold. By looking at the difference in intercepts between the regressions, one can evaluate the treatment effect. The results of this paper show a positive and significant treatment effect that is robust to the inclusion of the additional covariates book-to-market ratio, market value and SIC industry dummies. There is a positive and significant correlation between the market value and the stock return, but the other covariates are insignificant. Moreover, the McCrary sorting test has shown a possible manipulation of the forcing variable. Therefore, a donut RD design was also used. This removed the possibly manipulated observations. With the removal of the observation, the treatment effect becomes insignificant.

The paper is structured as follows: Section 2 discusses the existing literature. Section 3 is about the data, and section 4 explains the methodology. Section 5 presents the results, and finally, Section 6 is the conclusion. Appendix A includes additional estimation details, Appendix B contains additional tables, and Appendix C is a simulation study to show the difference between regular and donut RD design estimates.

2 Literature

This research links to several different pieces of literature. This section reviews existing literature on the topic. As well as literature on behavioural biases that can explain investors' behaviour.

There is existing research that shows that US firms exhibit higher returns during the period when earnings are announced. For example, Frazzini and Lamont (2007) showed that earnings announcements lead to higher returns before, on the day of the earnings announcement and after. According to their research, the higher returns on the day of the announcement and in the period thereafter can be explained by the volume hypothesis. This means that with an increase in trading volume, there are higher returns. They suggest that this increase in volume may come from the fact that more individual investors buy stocks with earnings announcements because the announcement of earnings brings more attention to a company. They found that this premium is positively correlated with a firm's market capitalization. An explanation for this could be that there is more news about large stocks. A trading strategy wherein a long position is taken in stock that will announce earnings in the coming month and a short position is taken in stocks that are expected not to generate a 60 basis points monthly return. Their research does not cover whether the earnings actually meet the expectations, and they only look into monthly returns.

Barber et al. (2013) found that this premium is not exclusive to the US stock market. They studied the premiums for 20 different countries, and for nine of them, they found similar results. They perform a linear regression with returns as the dependent variable and an indicator variable to indicate whether a firm is about to announce its earnings, as well as the additional explanatory variables book-to-market ratio, market value, and country fixed effects. Their results show that the coefficient for the indicator function is positive and significant and that returns are, on average, 1% higher during months of earnings announcements. They also show that the results are robust over time. Contrary to Frazzini and Lamont (2007), they found that smaller firms exhibit higher returns in the global sample. They checked if the higher returns resulted from unexpectedly strong earnings during the sample period, but this was not the case. This result aligns with the "small firm effect", where small firms have higher returns in general (van Dijk, 2011). A possible explanation for this effect is the risk increase linked to smaller firms where liquidity is a significant risk factor (Crain, 2011). Next to firm size, firm growth also plays a role. Growth stocks react more heavily to negative earnings surprises than positive earnings surprises (Skinner and Sloan, 2002).

Insider trading can mitigate or strengthen the drift. After an earnings announcement, investors are not sure if there is a structural change in earnings when a considerable earnings surprise occurs. Insiders have more information about this, and therefore, their trading patterns provide valuable information. This can mitigate the drift when insiders trade opposite the market and strengthen the drift when insiders confirm the market direction (Dargenidou et al., 2018).

Yan et al. (2020) studied the relation between the immediate and subsequent effects of earnings surprises. Their results show that when positive (negative) earnings are announced, the drift is larger (more negative) for firms with a smaller immediate effect. This implies that the larger the immediate effect is, the smaller the drift. Firms with the smallest immediate market response show a drift of 2.8% over three months, whereas firms with the most prominent immediate market response only show a drift of 1.4%. Firms with a larger market cap show higher immediate responses and their earnings surprises are more minor, and the returns are larger. The observed drift goes against the efficient market hypothesis, which states that all available information is incorporated in share prices (Berk and Demarzo, 2019). If this holds, the share price should immediately reflect new information. The drift implies that past information is still affecting current stock returns. Firms that outperform the expectations move up for a more extended period, and stocks that underperform move down.

Several behavioural biases can offer an explanation for the observed drift, one of which is overconfidence. In general, people are overconfident (Agner, 2020). This means that investors tend to overestimate their abilities. Under this bias, a positive relationship exists between prior returns and trading volume. Research on the Taiwanese stock market has shown that institutional investors exhibit this too, although to a lesser extent than individual investors (Chou and Wang, 2011). Another bias is the self-attribution bias, where people think that successful outcomes are because of their skills and adverse outcomes are bad luck Hoffmann and Post (2014). A combination of these biases causes investors to assign less value to public information than their own analysis, which may lead to an underreaction following an earnings announcement (Daniel et al., 2002). This can cause a drift in the stock returns instead of an instantaneous increase in the share price.

The literature discussed above shows existing research on the longer-term effect of earnings surprises on stock returns. This research shows the presence of an upward (downward) drift in case of a positive (negative) earnings surprise. Current research lacks the immediate effect of earnings surprises.

A behavioral theory that constitutes the relevance of looking at the immediate effect is the overreaction hypothesis. De Bondt and Thaler (1985) showed that, in the stock market, people tend to overreact to unanticipated events. They compared portfolios consisting of prior winners and prior losers. The latter had 25% higher returns over 36 months. This suggests price corrections back to the actual value of the stock following prior overreactions that caused the stock price to decrease too much. With the presence of overreacting investors, significant price movements following unanticipated earnings surprises can be expected. This price movement is captured by looking at the immediate returns following the announcement. The confirmation bias can strengthen the overreaction. According to this theory, people value information that confirms their beliefs more heavily than other information (Agner, 2020). In case of a positive earnings announcement, investors with a long position, who believe that it is a good company, will value this information more heavily than investors with a short position for whom the information contradicts their beliefs. This could cause a net overbuying of the stock. For negative earnings announcements, the same reasoning can be used oppositely.

Earnings are typically announced during after-hours trading after the market has closed. When looking into the immediate effect of earnings surprises, most of the price movement happens during this period and not on the actual trading day thereafter, as shown by Barclay and Terrence (2015). After the earnings announcement, there is a large spread between the bid and ask prices. Following a positive earnings surprise, the ask prices adjust quickly, followed by a slower-reacting bid price. For negative news, the reverse is true, as shown by Bazinas (2021). Moreover, they found that negative surprises have a more significant impact than positive surprises, and the changes in share price are caused mainly by liquidity providers updating their prices instead of actual trading. This is shown by graphing the trade and midquote, i.e. the mean between the bid and ask price. There is a quicker adjustment in midquote returns compared to trading returns for S&P 500 stocks. Even though these papers look at the immediate effect of an earnings announcement, only the trading patterns are considered, not the actual returns.

3 Data

The data consists of analysts' expectations, actual earnings, stock returns, S&P 500 returns and several firm-specific variables of all the companies currently in the S&P 500. The firm-specific variables are the book-to-market ratio, the market capitalization and the firm's industry. Market capitalization measures the firm's size and is the sum of all issue-level market values, measured in billions of dollars. The book-to-market ratio measures growth and is calculated by dividing the market capitalization by the common shareholder equity. Standard Industrial Classification (SIC) is used for the industry classification. This places firms into four-digit industry codes. The first digit represents the broad sector, the second digit is a more narrowed-down major group, the third is the industry group, and finally, the fourth digit represents the exact industry. These values for the firm-specific variables are measured on the last day of each quarter.

For each company, several years of data were used. The period runs from January 2000 to December 2023. Each year, there are 4 data points for each company, namely, one for every quarter. Not all companies have data over the entire period; the maximum available data was used for these companies. The data was collected from Wharton Research Data Services (WRDS). The analysts' expectations and actual earnings come from the Institutional Brokers' Estimate System (I/B/E/S). Multiple analysts make forecasts for the same quarter of the same company. To account for this, an average of all available expectations for each company is taken for every period. The stock returns and the S&P 500 returns come from the Center for Research in Security Prices (CRSP) and the firm-specific data comes from Compustat. Stock returns are calculated as the percentage change between the previous day's closing price and the current day's closing price. Calculating it this way ensures that the returns caused by after-hours and pre-market trading are also included.

Some observations lack data on a single or multiple variables. These observations have been removed. In total, there are 26,290 complete observations.

In the research, the surprise ratio is used to measure the size of the earnings surprise, and the excess return over the S&P 500 index serves as the dependent variable. The surprise ratio is calculated as follows:

$$\mathrm{SR}_{it} = \frac{\text{actual earnings}_{it}}{\text{expected earnings}_{it}} - 1. \tag{1}$$

Where SR_{it} is the surprise ratio for firm i in period t. The excess return is defined as the difference between the return of a stock and the return of the S&P 500 on the same day. The descriptive statistics for the variables are shown in Table 1. The surprise ratio shows large outliers, but only a small bandwidth around the threshold is considered in this research so this will not be an issue.

Table 1: Descriptive statistics

Variable	Mean	Std. dev.	min	max
Surprise Ratio	0.11	10.27	−594.24	1 385.50
Excess return	0.00	0.04	−0.40	0.52
Book-To-Market ratio	0.40	0.61	−0.40	0.52
Market value	45.06	111.14	0.09	3 035.22

Note. This table shows several descriptive statistics for the variables used.

4 Research methodology

4.1 Regression Discontinuity Design

The causal effect of a treatment can be analyzed through regression discontinuity (RD) design. This method is used when there is a variable that determines the assignment to a specific treatment. This variable is called a forcing variable. When the forcing variable exceeds a certain threshold, the treatment is applied. This paper uses the surprise ratio (SR) as the forcing variable in the RD design. In this case, the "treatment" means that the actual earnings are higher than expected. This is true when the surprise ratio exceeds zero.

There are two types of RD designs: sharp and fuzzy. In a sharp RD design assignment to the treatment group happens exactly at the threshold. Moreover, in a fuzzy RD design, treatment assignment is not perfectly determined by a variable crossing the threshold. The method used is the sharp RD since there is a fixed threshold, namely, zero. Let $Y_{it}(1)$ be the excess return over the S&P 500 index for firm i in period t in case the treatment is applied and $Y_{it}(0)$ otherwise. Then the treatment effect τ can be defined as follows:

$$\tau = \mathbb{E}[Y_{it}(1) - Y_{it}(0)|\text{SR} = 0]. \tag{2}$$

Two assumptions have been made. The most important assumption is that the only thing distinguishing observations on either side of the cutoff is the assigned treatment. This means that there is no manipulation of the forcing variable around the cutoff. It is assumed that $F_{Y(0)|X}(y|x)$ and $F_{Y(1)|X}(y|x)$ are continuous in x for all y. This means that the distributions of outcomes should change smoothly. Under these assumptions

$$\mathbb{E}[Y(0)|SR = 0] = \lim_{x \uparrow 0} \mathbb{E}[Y(0)|SR = x] \tag{3}$$

and

$$\mathbb{E}[Y(1)|SR = 0] = \lim_{x \downarrow 0} \mathbb{E}[Y(0)|SR = x]. \tag{4}$$

The treatment effect can then be defined as follows:

$$\tau = \lim_{x \downarrow 0} E[Y_{it}|\text{SR}_{it} = x] - \lim_{x \uparrow 0} E[Y_{it}|\text{SR}_{it} = x]. \tag{5}$$

The treatment effect is estimated with local polynomial regressions that are discussed in Section 4.2.

4.2 Local Polynomial Regressions

To estimate the conditional expectations in Equation (5) local polynomial regressions around the threshold have been fitted. The same regression is fitted on either side of the threshold by means of weighted least squares. Weights are determined based on a kernel function where non-zero weights are given to variables within a certain bandwidth with forcing variable surprise ratio$_{it}$ $\in [-h, h]$. The method for determining this bandwidth is discussed in Appendix A.3. Within the bandwidth, the weights are determined based on a kernel function. As suggested by Imbens and Lemieux (2008), a rectangular kernel function is used as they argue that there not much difference when a more sophisticated kernel is used. This means that these are regular regressions. The following general regression is used in the paper for values below the threshold:

$$Y_{it} = \alpha^- + \sum_{k=1}^{p} \beta_k^- (\text{SR}_{it})^k + \sum_{l=1}^{q} \gamma_l (Z_l)_{it} + \epsilon_{it} \quad \text{for} \quad -h < \text{SR}_{it} < 0. \tag{6}$$

For values above the threshold, the following regression is used:

$$Y_{it} = \alpha^+ + \sum_{k=1}^{p} \beta_k^+ (\text{SR}_{it})^k + \sum_{l=1}^{q} \gamma_l (Z_l)_{it} + \epsilon_{it} \quad \text{for} \quad 0 < \text{SR}_{it} < h. \tag{7}$$

In Equations (6) and (7), Z_l represents an additional covariate. And ϵ_{it} represents the error term. The coefficients for the covariates are assumed to be the same above and below the threshold to ensure that any observed discontinuity at the threshold is only caused by the treatment effect. The - (+) superscript indicates that this coefficient corresponds to the regression below (above) the threshold. For the first and second sub-questions, the covariates are excluded from the regression. For the third subquestion, the covariates market cap, book-to-market ratio and industry are included. As mentioned in the data section, the SIC industry codes are used for the industry. As this industry classification system uses four digits there are many different industries. For the dataset that was used there are 180 different industries. To limit the number of dummy variables only dummy variables are added for the first digit of the SIC code. To prevent perfect multicollinearity the dummy variable for the SIC codes starting with nine is dropped. In the dataset, no SIC codes are starting with zero. This means that there are eight industry dummies. Estimates are obtained by solving the minimization problem as shown in Appendix A.2.

$\hat{\alpha}^+$ is the estimate for $\lim_{x \uparrow 0} E[Y_{it}|\text{SR}_{it} = x]$ and $\hat{\alpha}^-$ is the estimate for $\lim_{x \uparrow 0} E[Y_{it}|\text{SR}_{it} = x]$ the treatment effect is now estimated as,

$$\hat{\tau} = \hat{\alpha}^+ - \hat{\alpha}^-. \tag{8}$$

5 Results

5.1 Bandwidth

In Table 3, the optimal bandwidths are shown for different values of δ for different polynomial orders. Also, the bandwidths are shown for models with and without covariates. A smaller δ makes the expected value of the cross-validation criterion closer to the actual value but increases noise as the bandwidth becomes larger. The bandwidths are larger for the models with covariates. This can be explained by the increased complexity of the model. To account for the additional variables, a larger bandwidth may be necessary. Following Imbens and Lemieux (2008), we let $\delta = 0.5$. This strikes a balance between bias and noise. Moreover, the same δ is used across all different models as well as for the specifications tests.

5.2 Regression discontinuity design results

Table 2 presents the results for the local polynomial regressions for the model without additional covariates. The bandwidths for these regressions are the bandwidths from the previous section. Four different polynomial orders are considered. Across the different models, the estimated treatment effect is similar. For the models with polynomial orders three and four, the treatment effect is not significantly different from zero at a significance level of 5%. The linear model without additional polynomial terms estimates the largest treatment effect. In this model, the treatment effect is equal to 0.003. This means that when a firm receives the treatment, i.e. has a positive earnings surprise, the stock return on the day following the earnings announcement is, ceteris paribus, 0.3% higher compared to the returns without receiving the treatment.

To see how well the models perform against each other, the adjusted R-squared and root mean square error (RMSE) are considered. The model with polynomial order one has the lowest adjusted R-squared. In terms of RMSE, it has the best performance. The inclusion of a second-order polynomial term improves the adjusted R-squared significantly, and the coefficient for the second-order polynomial term is found to be significant above the threshold. Higher-order polynomial terms in the model do show an improvement in adjusted R-squared, but the coefficients are not significant.

In all models, the constants are negative, both below and above the threshold. Despite the constants being negative, the constant above the threshold is higher than the constant below the threshold. This implies a positive treatment effect, which aligns with prospect theory, where the discontinuity is caused by asymmetrical weighting. The coefficients for β_1 are positive on both sides of the threshold, implying that the surprise ratios and excess returns are positively correlated. One would also expect this, as higher earnings surprises should be considered good news. For positive earnings surprises, the size of the coefficient is more than twice as high as for negative earnings surprises. This contradicts prospect theory. It may be true that the theory still holds, but that other biases in investor behaviour are way more heavily. In this case, confirmation bias might play a role. When investors invest in a company, they do so because they think that it is a good company. When news that confirms their beliefs is released, they value this more than bad news that contradicts their beliefs. Looking at the β_2 coefficients, there is no significant curvature found on the negative side of the threshold. In the models with polynomial orders two and three, the coefficient for β_2 is negative and significant. This implies that the curve is concave and that there is a decline in the effect of the earnings surprise on the stock returns for larger values of the surprise ratio. This aligns with an important economic concept, the law of diminishing marginal utility. According to this law, a person's marginal utility curve is concave Frank and Cartwright (2013). In this case, the positive effect of an additional increase in the surprise ratio decreases as the surprise ratio is larger. None of the coefficients corresponding to the higher-order polynomial terms are found to be significant.

Table 2: Regressions table without additional covariates for different polynomial orders

Variable \ Order	1	2	3	4
α^-	−0.004	−0.004	−0.004	−0.004
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
α^+	−0.002	−0.002	−0.002	−0.002
	(0.001)**	(0.001)**	(0.001)**	(0.001)**
β_1^-	0.030	0.039	0.056	0.026
	(0.014)*	(0.019)*	(0.023)*	(0.034)
β_1^+	0.071	0.092	0.092	0.097
	(0.008)***	(0.011)***	(0.014)***	(0.022)***
β_2^-		0.071	0.180	−0.095
		(0.060)	(0.113)	(0.251)
β_2^+		−0.160	−0.181	−0.223
		(0.037)***	(0.069)**	(0.160)
β_3^-			0.146	−0.594
			(0.139)	(0.610)
β_3^+			0.104	0.190
			(0.086)	(0.394)
β_4^-				−0.589
				(0.459)
β_4^+				−0.045
				(0.303)
Bandwidth(h)	0.185	0.516	0.618	0.764
Cutoff(c)	0.000	0.000	0.000	0.000
Treatment effect(τ)	0.003	0.002	0.002	0.002
	(0.001)**	(0.001)*	(0.001)	(0.001)
Adjusted R-Squared	0.013	0.021	0.023	0.023
RMSE	0.038	0.040	0.041	0.041

Note. This table shows the regression coefficients and the corresponding treatment effect for the case without additional covariates. The dependent variable is the excess return. Standard errors are between parenthesis. Significance levels: '***' 0.001 '**' 0.01 '*' 0.05

5.3 Regressions with additional covariates

The results for the regressions with additional covariates are shown in Table 4. With the addition of extra covariates, the α coefficients become insignificant. The treatment effect remains positive and significant for the models with polynomial orders one and two. For these models, the treatment effect is twice as large as those without covariates. This may be caused by the additional covariates or the larger bandwidth that is now used. The β coefficients differ in magnitude compared to the models without covariates but have the same sign. The coefficient corresponding to the market value is significant for all polynomial orders. The effect is negative, implying that larger firms, in terms of market cap, experience lower returns on the day following the earnings announcement compared to smaller firms, though the effect is small. For the mean market value of 45 billion dollars, this would result in an effect of approximately 0.02%. This contradicts the results from Frazzini and Lamont (2007) but aligns with the results from Barber et al. (2013). All industry dummies and the book-to-market ratio are insignificant. The adjusted R-squared is slightly higher for the models with covariates but so are the RMSE's suggesting that the model's fit did not necessarily improve.

5.4 Specification tests

Three specification tests were performed to verify the observed results. The subsequent sections show the results of the tests. The treatment effect is tested to determine whether it is robust to the removal of potentially manipulated data points and to different bandwidth sizes. A threshold at other non-discontinuity points is evaluated to determine whether it does not show any treatment effect.

5.4.1 Donut RD

A McCrary sorting test, to test for continuity of the distribution of the forcing variable, rejects the null hypothesis of continuity even at a 1% significance level. The p-value for the test is equal to 0.00. Continuity of the forcing variable is not required for the regression discontinuity design, but it suggests that the forcing variable is potentially manipulated. This could mean that firms alter their earnings to beat expectations or that analysts report lower forecasts following the under promise over deliver approach. Firms can influence their earnings by using different accounting techniques. Myers et al. (2007) showed that firms indeed do this, but it is unclear whether or how many observations are affected by it. There may be other reasons why the null hypothesis is being rejected other than manipulation of the forcing variable. It could, for instance, also be random variation. To test if the treatment effect is still significant with the removal of the potentially manipulated observations donut RD is used. Appendix A.4 provides an explanation for the donut RD method, and Appendix C illustrates the benefit of donut RD with a simulation study.

Table 5 shows the treatment effect for a donut hole sizes equal to 0.025, 0.05 and 0.1. Again, the bandwidths were obtained using the cross-validation approach. The results show a significant treatment effect for the model with polynomial order one and the donut hole size equal to 0.1. There are no significant treatment effects for the other donut hole sizes or polynomial orders. These results do not rule out that the treatment effect is caused by manipulation of the forcing variable. The fact that the treatment effect is shown to be significant for a donut hole size equal to 0.1 could indicate that the treatment is robust to the removal of manipulated data. The insignificant treatment effects for the other donut hole sizes suggest that the initially found treatment effect may be caused by the manipulation of the forcing variable around the threshold. The exclusion of the observations in an interval around the threshold potentially removes relevant observations. With the removal of all observations within the interval [-0.05,0.05], more than 10,000 observations are removed. Therefore, these results can not guarantee that the treatment effect is only caused by manipulation. It is out of the scope of this research to analyze this in further detail, but future research should look into this.

5.4.2 Different bandwidths

There are different methods to compute the bandwidth. Moreover, the δ parameter can be changed in the cross-validation approach used in this paper. This affects the bandwidth. Table 6 shows the treatment effect for different bandwidths. The actual bandwidth is equal to h, and multiples of this bandwidth have been considered. The table shows that the treatment effect is different for different bandwidth sizes. It remains positive and significant in the cases where it was significant with the original bandwidth. For several models, a different bandwidth causes the treatment effect to become significant. There is no clear relation between the treatment effect and the bandwidth size. For some of the models, the treatment effect increases with the bandwidth size, but for example, the model without additional covariates and polynomial order four shows the largest treatment effect for the smallest bandwidth size. This shows that the treatment effect remains for larger and smaller bandwidth sizes.

5.4.3 Jumps at non-discontinuity points

Another specification test checks whether there is a discontinuity at other points where it is not expected. The same method is used as suggested by Imbens and Lemieux (2008). In their method, the data is split at the cutoff. So, in this case, there is a subsample with all observations with $SR < 0$ and another subsample with all observations with $SR \geq 0$. An RD design model is fitted for both subsamples where the cutoff is set equal to the median surprise ratio within that subsample. It is important to note that the actual discontinuity point is not in there by only using observations on either side of the threshold. The results for the test are shown in Table 7. For both subsamples, no significant treatment effects were found. This is also what was expected, as there is no treatment at these points.

6 Conclusion

This paper analyzed the effect of earnings surprises on the stock returns of the first trading day following the earnings announcement. A surprise ratio was constructed, serving as the forcing variable in the regression discontinuity design. Multiple polynomial orders have been considered as well as several additional covariates. The bandwidths were selected with a cross-validation approach. The results show that the models with polynomial order two provide the best fit. At the same time, higher-order polynomial terms are found to be insignificant. Both with and without the inclusion of additional covariates, a positive and significant treatment effect is present for the models with polynomial orders one and two. This implies that a positive earnings surprise results in higher returns than a negative one. It has been found that a firm's market value has a small but positive and significant effect on the stock returns following an earnings announcement. The book-to-market ratio and SIC industry dummies do not show any significant effects. Adding a second-order polynomial term improves the model's explanation power for the model with and without additional covariates. Moreover, the first-order polynomial term coefficients are positive on both sides of the threshold. This means that there is a positive correlation between the surprise ratio and the stock returns. Below the threshold, there is no significant curvature, while there is a concave relationship above the threshold. Therefore, the effect of a larger earnings surprise on the stock returns declines for higher values of the surprise ratio. The McCrary sorting test rejects the null hypothesis of continuity of the forcing variable at the threshold. This may mean that the forcing variable is manipulated. A donut RD design was performed, showing that removing observations around the threshold made the treatment effect disappear for all donut sizes and polynomial orders except the case with donut size equal to 0.1 and polynomial order one. The problem with this approach is that, potentially, many relevant observations have also been removed. Further research should analyze the reason for rejecting the continuity of the forcing variable at the threshold. This can show whether there is a manipulation of the forcing variable. Firms manipulating their earnings to beat analyst expectations can affect investor trust and overall trust in the financial market. Potentially, there should be more regulation concerning earnings manipulations.

Another suggestion is to use machine learning methods for the RD design. With machine learning methods, one can potentially include more covariates than the number of observations. Kreiss and Rothe (2022) have shown that this can improve the accuracy of the treatment effect. This paper only found that a firm's market value has a significant effect. Perhaps other industry classifications or financial ratios can be used. Another potentially significant covariate is a seasonality dummy, as in January, returns tend to be higher Moller and Zilca (2008). Machine learning can also determine the optimal bandwidth and polynomial order. Long and Rooklyn (2020) developed a method to reduce the MSE by over 20%.

References

Agner, E. (2020). *A Course in Behavioral Economics.* Bloomsbury Academic.

Barber, B. M., De George, E. T., Lehavy, R., and Trueman, B. (2013). The earnings announcement premium around the globe. *Journal of Financial Economics*, 108(1):118–138.

Barclay, M. J. and Terrence, H. (2015). Price discovery and trading after hours. *The Review of Financial Studies*, 16(4):1041–1073.

Bary, E. (2024). Spotify's stock rallies after earnings as user growth brings a positive surprise. *MarketWatch.* `https://www.marketwatch.com/story/spotifys-stock-rallies-after-earnings-as-user-growth-brings-a-positive-surprise-6ab07371`.

Bazinas, S. (2021). Earnings announcements retail participation: A deep-dive into after-hours trading. *NYSE.* `https://www.nyse.com/data-insights/earnings-announcements-and-retail-participation-a-deep-dive-into-after-hours-trading`.

Berk, J. and Demarzo, P. (2019). *Corporate Finance, Global Edition.* Pearson Education Limited.

Brown, L. D. (1997). Analyst forecasting errors: Additional evidence. *Financial Analysts Journal*, 53(6):81–88.

Chou, R. K. and Wang, Y.-Y. (2011). A test of the different implications of the overconfidence and disposition hypotheses. *Journal of Banking Finance*, 35(8):2037–2046.

Crain, M. A. (2011). A literature review of the size effect.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (2002). Investor psychology and security market under- and overreactions. *The Journal of Finance*, 53(6):1839–1885.

Dargenidou, C., Tonks, I., and Tsoligkas, F. (2018). Insider trading and the post-earnings announcement drift. *Journal of Business Finance Accounting*, 45(3-4, SI):482–508.

De Bondt, W. and Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, 40(3):793–805.

Frank, R. H. and Cartwright, E. (2013). *Microeconomics and Behaviour.* McGraw-Hill Education.

Frazzini, A. and Lamont, O. A. (2007). *The Earnings Announcement Premium and Trading Volume.* (NBER Working Paper No w13090). National Bureau of Economic Research. `https://www.nber.org/system/files/working_papers/w13090/w13090.pdf`.

Hoffmann, A. O. and Post, T. (2014). Self-attribution bias in consumer financial decision-making: How investment returns affect individuals' belief in skill. *Journal of Behavioral and Experimental Economics*, 52:23–28.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Kreiss, A. and Rothe, C. (2022). Inference in regression discontinuity designs with high-dimensional covariates. *The Econometrics Journal*, 26(2):105–123.

Long, M. C. and Rooklyn, J. (2020). Next: Machine learning for regression discontinuity. Technical report, Evans School of Public Policy and Governance.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714. The regression discontinuity design: Theory and applications.

Moller, N. and Zilca, S. (2008). The evolution of the january effect. *Journal of Banking Finance*, 32(3):447–457.

Myers, J. N., Myers, L. A., and Skinner, D. J. (2007). Earnings momentum and earnings management. *Journal of Accounting, Auditing & Finance*, 22(2):249–284.

Naseer, M. and bin Tariq, Y. (2016). The efficient market hypothesis: A critical review of the literature. *The IUP Journal of Financial Risk Management,*, 12:48–63.

Noack, C. and Rothe, C. (2023). Donut regression discontinuity designs.

Patell, J. M. and Wolfson, M. A. (1979). Anticipated information releases reflected in call option prices. *Journal of Accounting and Economics*, 1(2):117–140.

Ruffino, G. (2024). Tesla stock drops as earnings fall short of analyst estimates. *Euronews.* `https://www.euronews.com/business/2024/01/25/tesla-income-doubles-but-earnings-fall-short-of-analyst-estimates`,.

Saul, D. (2017). Disney stock limps to worst day since 2022 despite surprise disney+ profit. *Forbes.* `https://www.forbes.com/sites/dereksaul/2024/05/07/disney-stock-limps-to-worst-day-in-12-months-despite-surprise-disney-profit/?sh=7d2a11ce74af`.

Skiba, A. and Skiba, H. (2017). *Financial Behavior: Players, Services, Products, and Markets.* Oxford University Press.

Skinner, D. J. and Sloan, R. G. (2002). Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies*, 7:289—312.

Ţiţan, A. G. (2015). The efficient market hypothesis: Review of specialized literature and empirical research. *Procedia Economics and Finance*, 32:442–449. Emerging Markets Queries in Finance and Business 2014, EMQFB 2014, 24-25 October 2014, Bucharest, Romania.

van Dijk, M. A. (2011). Is size dead? a review of the size effect in equity returns. *Journal of Banking Finance*, 35(12):3263–3274.

Yan, Z., Zhao, Y., and Fang, M. (2020). Immediate and subsequent market responses to earnings announcements. *China accounting and Finance Review*, 22(2):35–53.

A Estimation details

A.1 Programming

All programming was done using R. For the regression discontinuity results the package rdrobust was used and for the simulation study this was the RDHonest package.

A.2 Minimization problem

$$\min_{\alpha^-,\{\beta_k^-\}_{k=1}^p,\alpha^+,\{\beta_k^+\}_{k=1}^p,\{\gamma_l\}_{l=1}^q} \sum_i \sum_t (\mathbf{1}_{\{-h<\mathrm{SR}_{it}<0\}}[Y_{it} - \alpha^- - \sum_{k=1}^p \beta_k^- (\mathrm{SR}_{it} - x)^k] + \mathbf{1}_{\{0<\mathrm{SR}_{it}<h\}}[Y_{it} - \alpha^+ - \sum_{k=1}^p \beta_k^+ (\mathrm{SR}_{it} - x)^k] - \sum_{l=1}^q \gamma_l (Z_l)_{it})^2. \tag{9}$$

A.3 Bandwidth selection

The bandwidth for both equations is selected based on a cross-validation approach as in Imbens and Lemieux (2008). The same bandwidth is used on both sides of the cutoff, which the authors also suggested. This may not be the optimal solution, but it improves simplicity and consistency. The bandwidth is chosen to minimise the overall bias of $\hat{\tau}$. Following the defintion in Equation (8) this can be achieved by jointly minimising the biases for $\hat{\alpha}^-$ and $\hat{\alpha}^+$.

$$\min 0.5(E[\lim_{x\uparrow 0} E[Y_{it}|\mathrm{SR}_{it} = x] - \hat{\alpha}^-] + E[\lim_{x\downarrow 0} E[Y_{it}|\mathrm{SR}_{it} = x] - \hat{\alpha}^+]). \tag{10}$$

As the focus lies on the estimation around the threshold, observations in the tails are removed for the bandwidth selection. This reduces potential noise and improves running times. Only observations inside the interval $[q_{X,\delta,-}, q_{X,1-\delta,+}]$ are considered. Here $q_{X,\delta,-}$ the δ quantile of the empirical distribution of all the surprise ratios below zero and $q_{X,1-\delta,+}$ is the $(1-\delta)$ quantile of all the surprise ratios above and including zero. The regression from Equation (6) or (7) is fitted for each SR within this interval. For a SR below zero, this is the regression from Equation (6) with data from inside of the interval $(SR-h, SR)$ and for a SR above, this is from Equation (7) with data from inside of the interval $(SR, SR+h)$. For the case with additional covariates, these equations have to be fitted simultaneously to ensure that the covariates have the same coefficients on both sides of the threshold. For a SR below the threshold, the bias for $\hat{\alpha}^-$ can be calculated as $(Y_{it} - \hat{\alpha}^-)$ and for a SR above the threshold the bias for $\hat{\alpha}^+$ can be calculated as $(Y_{it} - \hat{\alpha}^+)$. This leads to the following definition of the cross-selection criterion:

$$CV(h) = \frac{1}{N} \sum_{i=1}^n \sum_{t=1}^m \mathbf{1}_{\{q_{X,\delta,-} \leq SR_{it} \leq q_{X,1-\delta,+}\}} (Y_{it} - \hat{\mu}(X_{it}))^2. \tag{11}$$

Where $\hat{\mu}(x)$ is defined as:

$$\hat{\mu}(x) = \begin{cases} \hat{\alpha}^-(x) \text{ if } x < c \\ \hat{\alpha}^+(x) \text{ if } x \geq c \end{cases}. \tag{12}$$

$\hat{\alpha}^-(x)$ and $\hat{\alpha}^+(x)$ are the solution of α^- and α^+ for following minimisation problem:

$$\min_{\alpha^-,\{\beta_k^-\}_{k=1}^p,\alpha^+,\{\beta_k^+\}_{k=1}^p,\{\gamma_l\}_{l=1}^q} \sum_i \sum_t (\mathbf{1}_{\{x-h<\mathrm{SR}_{it}<x\}}[Y_{it} - \alpha^- - \sum_{k=1}^p \beta_k^-(\mathrm{SR}_{it} - x)^k] + \mathbf{1}_{\{x<\mathrm{SR}_{it}<x+h\}}[Y_{it} - \alpha^+ - \sum_{k=1}^p \beta_k^+(\mathrm{SR}_{it} - x)^k] - \sum_{l=1}^q \gamma_l (Z_l)_{it})^2. \tag{13}$$

The optimal bandwidth now follows from:

$$h = arg\min_h CV(h). \tag{14}$$

A.4 Donut RD

In donut RD, design observations are removed at a small interval around the threshold. Further, all computations from Section 4 remain the same. The primary reason for removing these observations is that there may be sorting or manipulation of data around the threshold. This is tested using the density test from McCrary (2008). The test's null hypothesis states that there is no discontinuity in the density of the forcing variable at the threshold. Even if sorting or manipulation is present in the data, conventional RD may still provide better estimations. Section C presents a simulation study to show a comparison between conventional and donut RD estimates for different degrees of sorting.

B Additional tables

Table 3: Bandwidths for Different Deltas and Polynomial Orders with and without Covariates

	Without covariates			With covariates		
Order\\δ	0.5	0.7	0.9	0.5	0.7	0.9
1	0.145	0.185	0.094	0.522	0.554	0.235
2	0.378	0.647	0.236	0.880	0.880	0.596
3	0.656	0.752	0.434	0.702	0.609	0.685
4	0.764	0.931	0.634	0.807	0.260	0.428

Note. This table shows the different bandwidths for different deltas and polynomial orders for the cases with and without additional covariates.

Table 4: Regressions table with additional covariates for different polynomial orders

Variable \ Order	1	2	3	4
α^-	−0.006	−0.006	−0.005	−0.005
	(0.004)	(0.004)	(0.004)	(0.004)
α^+	−0.001	−0.002	−0.003	−0.003
	(0.004)	(0.004)	(0.004)	(0.004)
β_1^-	0.012	0.025	0.047	0.040
	(0.005)*	(0.009)**	(0.022)*	(0.033)
β_1^+	0.038	0.060	0.097	0.091
	(0.003)***	(0.006)***	(0.014)***	(0.021)***
β_2^-		0.040	0.125	0.064
		(0.014)**	(0.098)	(0.233)
β_2^+		−0.056	−0.211	−0.162
		(0.009)***	(0.062)***	(0.144)
β_3^-			0.069	−0.118
			(0.113)	(0.546)
β_3^+			0.148	0.010
			(0.071)*	(0.336)
β_4^-				−0.179
				(0.396)
β_4^+				0.116
				(0.244)
Book-to-market ratio	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Market value	−0.000	−0.000	−0.000	−0.000
	(0.000)**	(0.000)**	(0.000)*	(0.000)**
SIC1	0.002	0.001	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)
SIC2	0.001	0.001	0.000	0.000
	(0.004)	(0.004)	(0.004)	(0.004)
SIC3	0.001	0.001	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)
SIC4	0.001	0.000	0.000	0.000
	(0.004)	(0.004)	(0.004)	(0.004)
SIC5	0.002	0.002	0.002	0.002
	(0.004)	(0.004)	(0.004)	(0.004)
SIC6	0.000	−0.000	−0.000	−0.000
	(0.004)	(0.004)	(0.004)	(0.004)
SIC7	0.004	0.004	0.003	0.003
	(0.004)	(0.004)	(0.004)	(0.004)
SIC8	0.002	0.001	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)
Bandwidth(h)	0.522	0.880	0.702	0.807
Cutoff(c)	0.000	0.000	0.000	0.000
Treatment effect(τ)	0.006	0.004	0.002	0.002
	(0.001)***	(0.001)*	(0.001)	(0.001)
Adjusted R-Squared	0.022	0.024	0.023	0.024
RMSE	0.041	0.042	0.041	0.042

Note. This table shows the regression coefficients and the corresponding treatment effect for the case with additional covariates. The dependent variable is the excess return. Standard errors are between parenthesis. Significance levels: '***' 0.001 '**' 0.01 '*' 0.05 '

Table 5: Treatment effect for donut RD

Donut hole size \ Order	1	2	3	4
0.100	0.018	0.016	0.006	0.003
	(0.004)***	(0.010)	(0.010)	(0.015)
0.050	−0.001	0.002	−0.001	−0.006
	(0.005)	(0.003)	(0.004)	(0.005)
0.025	0.002	0.000	0.000	−0.000
	(0.002)	(0.003)	(0.002)	(0.003)

Note. This table shows the treatment effect for the donut regression discontinuity design for different donut hole sizes. The Standard errors are between parenthesis. Significance levels: '***' 0.001 '**' 0.01 '*' 0.05 '

Table 6: Treatment effect for different bandwidths

	Without covariates				With covariates			
Bandwidth\order	1	2	3	4	1	2	3	4
2h	0.004	0.004	0.003	0.002	0.008	0.006	0.003	0.002
	(0.001)***	(0.001)***	(0.001)***	(0.001)*	(0.001)***	(0.001)***	(0.001)**	(0.001)*
h	0.003	0.002	0.002	0.002	0.006	0.004	0.002	0.002
	(0.001)**	(0.001)*	(0.001)	(0.001)	(0.001)***	(0.001)*	(0.001)	(0.001)
0.5h	0.003	0.003	0.003	0.004	0.003	0.002	0.002	0.004
	(0.001)**	(0.001)*	(0.001)*	(0.001)*	(0.001)***	(0.001)*	(0.001)	(0.001)*

Note. This table shows the treatment effect for different bandwidth sizes. The actual bandwidth is equal to h. The Standard errors are between parenthesis. Significance levels: '***' 0.001 '**' 0.01 '*' 0.05 '

Table 7: Treatment effect for different points

	Without covariates				With covariates			
Subsample\order	1	2	3	4	1	2	3	4
$SR < 0$	0.001	0.000	0.001	0.000	0.003	0.001	0.001	0.000
	(0.002)	(0.003)	0.004	(0.005)	(0.002)	(0.003)	(0.004)	(0.005)
$SR \geq 0$	0.002	0.003	0.002	0.002	0.002	0.003	0.001	0.002
	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)	(0.003)

Note. This table shows the treatment effect for different, non-discontinuity, points. The Standard errors are between parenthesis. Significance levels: '***' 0.001 '**' 0.01 '*' 0.05 '

C Simulation

A simulation study was conducted to compare the regular RD design with the donut RD design. This is the same simulation as in Noack and Rothe (2023). The idea behind the simulation is to simulate both data where donut RD is not necessary and data where it is necessary. Then, both regular and donut RD are performed to compare the results. The forcing variable X_i is generated from a uniform distribution between -1 and 1. The outcome Y_i is defined as follows:

$$Y_i = \mu_L(X_i) + \epsilon_i, \tag{15}$$

with

$$\mu_L(x) = \text{sign}(x)x^2 - L\text{sign}(x)((x - 0.1\text{sign}(x))^2 - 0.1^2\text{sign}(x))\mathbf{1}_{|x|<0.1}. \tag{16}$$

For $L \in \{0, 10, \cdots, 40\}$ and $\epsilon_i \sim N(0, 0.5)$ independent of X_i. $\mu^*(x)$ is defined as $\mu_0(x)$. This means that $\tau^* = 0$. The definition of μ_L shows that L has an effect on the conditional expectation of the observed data compared to the hypothetical data over the area $(-0.1, 0.1)$. Larger values of L imply that a donut RD design is needed. Then, both conventional and donut RD estimates are computed using the package RDHonest with M=2. The donut hole size is equal to 0.1. This means that observations in the interval $[-0.1, 0.1]$ are excluded. The sample size in each simulation is equal to $n = 1000$. The simulation is repeated 10,000 times.

Tables 8 and 9 present the results of the simulation. The results differ from what Noack and Rothe (2023) have obtained. This may be caused by a difference in the version of the RDHonest package.

The bias, standard deviation and root mean square error (RMSE) for both the regular and donut RD estimators for different values of L are shown in Table 8. The table shows that when L increases, the bias and RMSE of the regular estimator increase in absolute value, whereas they remain constant for the donut estimator. The standard deviation is constant for different values of L for both RD estimators and is higher for the donut estimator. In terms of RMSE the donut estimator is better than the regular estimator for the cases with L larger than 10.

Next to the point estimations, it is also essential to look at the confidence intervals of the estimators. Table 9 shows the 95% confidence interval coverage and length. The table illustrates that the coverage for the regular RD intervals is not correct for larger values of L. However, this remains correct for the donut estimator. The length of the confidence interval is larger for the donut estimator, so for L=0, the regular estimator is preferred.

The results show that in the case of no sorting, regular RD is superior to donut RD, in the case where L=10 the regular RD has a lower absolute bias, standard deviation and RMSE. The coverage confidence interval is not entirely correct. Moreover, for larger values of L, the donut RD provides better estimates. This shows that for cases with minor sorting the regular method should still be considered.

Table 8: Bias, Standard Deviation, and RMSE for Regular and Donut Models

	Bias		Std. Dev		RMSE	
L	Regular	Donut	Regular	Donut	Regular	Donut
0	-0.049	-0.128	0.098	0.151	0.110	0.197
10	-0.119	-0.133	0.100	0.150	0.155	0.200
20	-0.186	-0.127	0.099	0.150	0.211	0.196
30	-0.254	-0.125	0.102	0.150	0.273	0.195
40	-0.320	-0.124	0.103	0.151	0.336	0.196

Note. This table shows the bias, standard deviation, and RMSE of the treatment effect for both the regular and donut RD models.

Table 9: Coverage and Length for Regular and Donut Models

	Coverage		Length	
L	Regular	Donut	Regular	Donut
0	0.9506	0.9454	0.430	0.749
10	0.8299	0.9440	0.430	0.748
20	0.6104	0.9500	0.430	0.750
30	0.3543	0.9521	0.431	0.748
40	0.1603	0.9502	0.430	0.749

Note. This table shows the coverage and length of the 95% confidence interval for the treatment effect of the regular and donut RD models.