PREDICTING ONLINE BOOKS SALES FROM REVIEW SENTIMENTS:

The Differential Impacts of Expert and Online Customer Reviews on Amazon.com's Book Sales and the Moderating Effect of Book Category

Student: Trang Doan
Student number: 378313

Supervisor: Dr. Jay Lee

MA Media & Business
Erasmus School of History, Culture and Communication
Erasmus University Rotterdam

Master's thesis
26-6-2017

ABSTRACT

The past decades have seen the advent of online book retail and its explosive growth. As the popularity of online book retail continues to increase, so does the interest in understanding the predictors of online book sales. Accordingly, prior research has found evidence for the significant effects of traditional (I.e. expert reviews) and social earned media (i.e. online customer reviews) on online book sales. By literature review, the researcher pinpoints three main drawbacks of the extant literature. First, scholars have only been studying the effect of online customer reviews and expert reviews in isolation, regardless of the fact that they actually interact with each other while influencing online book sales in reality. Second, little research has been carried out to closely examine the impact of sentiment reviews on online book sales with sophisticated automated sentiment analysis technique. Third, the moderating role of book category was insofar neglected in scholarship. Drawing from these perspectives, this research sets out to investigate the differential impacts of traditional earned media (i.e. expert reviews) and social earned media (i.e. online customer reviews) on Amazon.com's book sales. Additionally, it seeks to understand the moderating effect of book category on the impact of the positive sentiment in online customer reviews and expert reviews on Amazon.com's book sales. The research is guided by the following research question: *To what extent do traditional (i.e. expert reviews) and social earned media (i.e. online customer reviews) predict Amazon.com's book sales?* To answer the research question, the researcher conducts a quantitative research which consists of two steps of analyses. First, she performs a sentiment analysis on 1,133 reviews, using the sentiment analysis software LIWC 2015 (Linguistic Inquiry and Word Count), in order to detect the percentage of positive and negative sentiment in their textual content. Next, she conducts regression tests and z-score tests for regression coefficients to assess the hypotheses. The research provides practical recommendations for online book retailers to tailor their online communication strategy and budget so as to enhance sales performance. It, likewise, produces some important theoretical implications for future research to investigate the interplay between online customer reviews and expert reviews in influencing online book sales.

Keywords: e-commerce, online book retail, Amazon.com, online customer reviews, expert reviews, online sales

Table of Contents

1. INTRODUCTION

1.1. E-commerce and online book retail

E-commerce has come a long way since its commencement in the 1970s (Wigand, 1997). In the beginning, the term was used to describe the execution process of electronic commercial transactions with the assistance of current prominent technologies such as Electronic Data Interchange (EDI) and Electronic Funds Transfer (EFT) (Killian et al., 1994). As the Internet became more widespread among the general public and was opened to commercial use in 1990s, online retail came into mainstream popularity. E-commerce since then has been defined as selling and purchasing activities of goods and services over the Internet through secure connections and electronic payment service (Laudon & Traver, 2008). According to ActivMedia Research, the popularity of online purchasing has been rising given that 74% of the US digital population report to have made at least one online purchase. As of 2015, e-commerce takes up more than 10% of all retail sales.

Books were one of the earliest physical products being sold over the Internet. As a matter of fact, the first online stores predominantly offered computer books (Clay, Krishnan & Wolff, 2001). By 1999 – less than a decade since their beginning – books were already the second largest retail segment – after computers – sold over the Internet (BCG, 2000). Online book sales increased from inherently nothing in 1995 to over two billion dollars in 2000 (Forrester, 2001). In 2002, their sales accounted for 7.5 – 10% of total book sales in the United States (Cader, 2001; American Booksellers Association, 2002). Nielsen (2016) predicts a continued growth in online book purchases in accordance with a three-percent increase in sales of traditional print books in 2015. Considering their immense development, books are – unsurprisingly – the most-studied online retail category (Clay et al., 2001).

Within the international online book retail market, Amazon.com is a dominant online bookseller. It is one of the first e-commerce companies to sell printed books online with an initial selection of more than one million titles back in July 1995 (Chevalier & Goolsbee, 2003; Mosendz, 2014). Today, the company controls 65% of all new online book units, both in prints and digital copies. As for printer books sold online alone, it has a predominant share of 64%, which is intrinsically aligned with the increase in online sale of new printed books (Mosendz, 2014). Due to its tremendous influence within the market and its monstrous variety of book titles, Amazon.com has been chosen as the case study of online bookseller numerous times in previous scholarship (e.g. Clay et al., 2001, Brynjolfsson, Hu & Smith, 2003; Chevalier & Mayzlin, 2006, Cui, Lui & Guo, 2012, Bao & Chang, 2014).

1.2. The effects of traditional and social earned media on online book sales

The rise of online book retail has inspired scholars to investigate many of its sales performance predictors, and a large body of literature suggests that earned media have a significant effect on online book sales (e.g. Chevalier & Mayzlin, 2006; Zhu & Zhang, 2010; Hu, Hok & Reddy, 2013; Lee, Lee & Shin, 2014; Bao & Chang, 2014). Earned media consist of both traditional Internet sources (e.g. expert reviews) and user-generated content on online retail website (e.g. online customer reviews). While one strand of research suggests that traditional media publicity can affect sales, another finds that it can also be predicted by online customer reviews (e.g. Chevalier & Mayzlin, 2006, Duan, Gu & Whinston, 2008; Zhu & Zhang, 2010). However, this pool of literature, though substantial in amount, either focuses on traditional or social earned media, but rarely both (Stephen & Galak, 2012). Nevertheless, no form of media is active in isolation (Hu et al., 2013). Scholars have shown that traditional and social earned media influence one another and their effects are intertwined. That prior research has failed to take this into consideration leads to an existing knowledge gap in how earned media activities are generated, interdependent and subsequently impact online product sales.

In addition, the current literature on earned media effects on online product sales is limited due to the lack of empirical studies on review sentiments (Archak, Ghose & Ipeirotis, 2011). Furthermore, these studies employ a limited, simple categorical classification (positive versus negative) of review sentiments (e.g. Turney, 2002; Das & Chen, 2007). Because their methods do not measure the specific degree to which each emotion is expressed in the reviews, their findings do not fully reflect reality as a text may connote more than one type of emotion. Likewise, although prior research has found strong evidence that online customer reviews affect sales, they only use the numeric review ratings and the volume of reviews in their empirical analysis, and fail to consider the possible effect of textual information in reviews on product sales. Little research has sought to test whether the sentiments expressed in online customer reviews have an economic impact (Archak et al., 2011; Hu et al., 2013). That is to say, to the best of the researcher's knowledge, no research has attempted to examine the effect of sentiments in traditional earned media content, and only a few studies have been dealing with sentiments of online customer reviews. For the most part, the dependence on summary statistics of earned media and the neglect of their sentiments in past research results in a deficiency in our understanding of the impact of earned media on online product sales. Moreover, the causal link between earned media and online product sales is still inconclusive because previous researchers have not considered that the effect of earned media can vary depending on product category. Zhu and Zhang (2010) suggest that product characteristics play a crucial role in determining their influence on sales. Given these points, it is essential to assess the moderating role of product category on the relationship between earned media and online product sales.

Guided by the above introductory discussion, this study formulates the research question as follows:

RQ: *To what extent do traditional (i.e. expert reviews) and social earned media (i.e. online customer reviews) predict Amazon.com's book sales?*

To answer this research question, the research draws on Friestad and Wright's (1949) persuasion knowledge model and takes a quantitative approach. For sampling, Python scripts are written in order to crawl unstructured web data on books and their (online customer and expert) reviews. The analysis in this study includes two phases. First, the researcher performs automated sentiment analysis to detect the percentage of positivity and negativity in textual content of 1,133 reviews, using the sentiment analysis software LIWC 2015 (Linguistic Inquiry and Word Count). Then, regression analyses and z-tests for regression coefficients are conducted on both an aggregated dataset and a disaggregated one to assess the hypotheses (see the Theoretical Framework chapter). The findings from both datasets are eventually synthesized to produce a conclusion.

1.3. Academic and societal relevance

This study extends the extant literature on the effect of earned media on online product sales in several distinctive ways. First, prior studies focus exclusively on either the effect of online customer reviews or that of expert reviews on online product sales, while this research quantitatively investigates on how both types of earned media – whose textual content may contain both positive and negative sentiments – jointly affect Amazon.com's book sales. As a consequence, its findings will provide a broader picture – that is closer to reality – of the interplay between different kinds of reviews in affecting online product sales. Second, by taking into account that the effect of online customer reviews and expert reviews may vary for different product types, the researcher is able to pay a more comprehensive look at the mechanism of how they influence online book sales. Third, unlike previous studies on the moderating role of product type which are conducted with an experimental design, the research performs statistical analyses on actual e-commerce data from Amazon.com. Hence, its research results are applicable to more situations.

Additionally, predicting real world phenomena based on textual content of online communication is a rapidly growing area within sentiment analysis research (Patak, Mane, & Srivastava & Contractor, 2007; Ganiz, Pottenger & Yang, 2007). Prior to the 1990s, sentiment analysis research was limited due to slow computer infrastructure and the sheer complexity of qualitatively breaking down and organizing natural language into psychological states (Gothberg, 2012). Likewise, computational sentiment analysis has just emerged recently after the digitization of research documents became possible (Pannebaker et al., 2015). In this regards, this research paper contributes to the field by employing the most recent sentiment analysis technologies (i.e. LIWC 2015). Instead of simply establishing a binary classification of sentiments (positive versus negative), the researcher uses LIWC to calculate the degree to which positive and negative emotions are present in the textual content, which offers higher accuracy and minimizes biases by qualitative coders.

With respect to the fast growth of online book retail and the significant effects that earned media may have on them, the research moreover generates practical implications for online book retailers. It sheds some light on the simultaneous effects of online customer reviews and expert reviews on online book sales, which in turn better informs marketing professionals of how to better harness their power. Particularly, the empirical findings allow online book retailers to evaluate whether and to what degree they should make use of the two kinds of online communication activities concurrently to improve their sales. Hence, the research insights inform these practitioners in making decisions about allocating budget to varying online communication efforts. Furthermore, these insights can also be applied to other similar online product categories.

2. THEORETICAL FRAMEWORK

2.1. Persuasion knowledge model

Friestad and Wright's (1994) persuasion knowledge serves as the theoretical foundation for researching the effects of online customer reviews and expert reviews on Amazon.com's book sales in this study. Friestad and Wright's (1994) persuasion knowledge model is one of the most influential persuasion theories in advertising, marketing and consumer behavior research over the past 30 years (Schrum, Liu, Nespoli & Lowrey, 2012). Unlike other popular persuasion theories (e.g. the theory of reasoned action and elaboration likelihood model) that originate from social psychology, this model is specifically marketing-oriented (Schrum et al., 2012). Although the model has not yet been applied to examine the effects of expert reviews on online product sales, there are studies that employ it to investigate the persuasiveness of online customer reviews. For instance, Bambauer-Sachse and Mangold (2013) study how customers' knowledge about online product reviews affects the effectiveness of these reviews in influencing their product evaluations. In the same manner, Kusumasondijaja, Shanka and Marchegiani (2012) examine the moderation role of reviewer's identity in the relationship between review valence (positive vs. negative) and review credibility. These earlier studies are evidence of that Friestad and Wright's (1994) is a useful theoretical model for this study.

The persuasion knowledge model explains how people cope with persuasion attempts by accumulating and using personal knowledge about them. Henceforth, for this research, it helps investigating how customers' perception of the evaluative texts impacts their attitudes toward the products as well as their reviewers (Friestad & Wright, 1994). To more specific, the persuasion knowledge model posits the customers ("targets") form a personal understanding about presented information in online customer reviews and expert reviews ("persuasion attempts") which are constructed by previous customers and expert reviewers respectively (Friestad & Wright, 1994) (Figure 1). In their theory, Friestad and Wright (1994) highlight the proactive role of the targets in interpreting persuasion attempts. The persuasion attempts are not limited to what the agents want to communicate, but also include the targets' understanding of how, when and why the agents have designed the delivered "message". Provided that reviewers can publish online reviews either with or without persuasive intentions (Sundaram, Mitra & Webster, 1998), the underlining two-way interaction within the persuasion attempts makes theoretical model broad and flexible enough to accommodate reviewers' varying motives for posting reviews online.

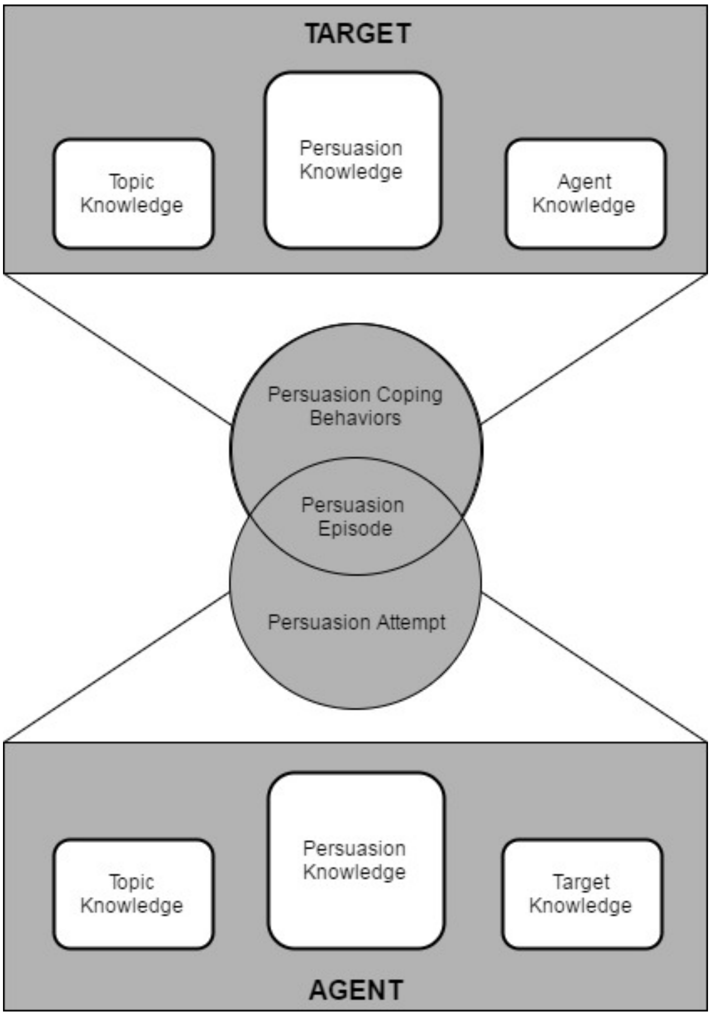


Figure 1. Friestad and Wright's (1994) persuasion knowledge model

Equally important, the persuasion knowledge model can direct the discussion about the moderating effect of book category on the impact that review sentiments have on online book sales. Since the model pays attentions to customers' active use of their contingent persuasion knowledge to cope with persuasion attempts, it helps form an understanding about how customers of experience products tend to consume more information embedded in online reviews when making a purchase compare to those of search products (Bei, Chen & Widdows, 2004). Hence, it provides a strong theoretical foundation to hypothesize the effects of different sentiments in online customer reviews and expert reviews on online book sales vary across books with more search good attributes and those with more experience good attributes.

2.2. Definition of online customer reviews and expert reviews

Both offline and online media activities can be summarized into three categories: paid media, owned media and earned media (Corcoran, 2009; Goodall, 2009). Within the scope of this research, earned media and their effects on online book sales are examined. Earned media refers to media activities that are generated by external stakeholders such as customers, in the case of word-of-mouth communication, or journalists, in the case of traditional media content. Marketing efforts by companies may result in earned media activities, but they do not directly produce them (Stephen & Galak, 2012). Liu (2006) finds evidence that consumers perceive these third-party sources of information as credible. Earned media can be subsequently divided into traditional and social earned media. The former covers published or broadcasted content on professional media outlets. The latter refer to online consumer-generated media content such as online customer reviews, blog posts, forum conversations or social media updates (Stephen & Galak, 2012).

This research centers on expert reviews (traditional earned media) and online customer reviews on Amazon.com (social earned media). There seems to be no literature that squarely discusses the definition of expert book reviews. Therefore, in order to define them, the researcher draws on previous literature on movie reviews by film critics (e.g. Eliashberg & Shugan, 1997; Holbrook, 1999; Basuroy, Chatterjee & Ravid, 2003) given that books and movies harbor substantial similarity in terms of their experiential nature. The content of expert reviews has a mixture of professional critique and personal opinions written by book critics (Holbrook, 1999). These writings are, however, more likely to concentrate on the technical or artistic aspects of the book and their creators tend to showcase their expertise (Holbrook, 1999). Most of the time, recognized newspapers websites are the place to find these reviews.

Meanwhile, online customer reviews are feedback by customers who previously purchased a specific product via an online retail shop and they are posted directly onto e-commerce sites (Cui et al., 2012). The reviews consist of a textual evaluation of the purchased product and rating which offers a numerical score for its overall perceived quality (Amblee & Bui, 2007). For instance, Amazon.com lets customers to post reviews, rate the product (one scale 1 to 5, with 5 stars being the highest) and vote – either positively or negatively – on other customers' reviews in regards to their perceived helpfulness (Chen, Dhanasobhon & Smith, 2008). Although online customer reviews are criticized for being biased towards either extremely positive or negative sentiments, they offer highly valued opinions from laypeople's perspectives and thus have become an important factor in purchasing decisions for web-savvy consumers (e.g. Chevalier & Mayzlin, 2003; Duan et al., 2005; Amblee & Bui, 2007).

2.3. Effects of online customer reviews and expert reviews on book sales

2.3.1. Effects of the percentage of positive and negative sentiment in online customer reviews on book sales

In recent years, there has been a growing interest in understanding the effects of social earned media on product sales (Stephen & Galak, 2012). However, Archak, Ghose and Ipeirotis (2011) observe that available studies on this topic have only used numeric ratings, volume or length of reviews. Yet, there is only a handful of studies that examine the information embedded in the textual content. A commonly used metric used to analyze information contained in the review text is valence, defined as “the fraction of positive and negative opinions, carries important information about a product’s quality and serves as a recommendation for consumers” (Cui et al., 2012, p. 39). Positive information indicates the high quality and reputation of a product, whereas, negative information demonstrates dissatisfaction with the product (Cui et al., 2012). The overall sentiment of the textual content can significantly affect customer perceptions and purchase decisions (Liu, 2006; Miller, Fabian & Lin, 2009; Archak et al., 2011), given that customers read review text rather than solely relying on ratings (Chevalier & Mayzlin, 2006; Mudambi & Schuff, 2010). This is confirmed by Hu et al.’s (2013) panel research whose findings suggests that ratings do not have a significant, direct effect on Amazon.com’s book sales, but rather sentiments expressed in the written content do. In their research, 58% of the respondents perceive numerical ratings in the initial phase of book search and awareness, but its importance decreases as they shift into the purchase stage. In the purchase stage, 65% of the respondents consider text sentiments as an important cue for their purchase decision and its importance grows during this phase (Hu et al., 2013).

The literature on impression formation proposes that, when comparing negative with positive information, people have a tendency to place more importance on negative information while assessing a product (Fiske, 1980; Skowronski & Carlston, 1989). As explained by Cui et al. (2012), this phenomenon is widely explained by the negativity bias, a psychological tendency to assign more weight to negative information during product evaluation. Zhang, Cranciun and Shin (2011) examine the persuasiveness of online customer reviews and assert that customers, when evaluating products associated with promotion consumption goals (e.g. books), are more likely to be convinced by negative reviews. That is, negative reviews have a negative effect on product sales. Cui et al. (2012) confirm this finding and also observe that positive online customer reviews increase sales of experience products (e.g. books).[1] Furthermore, they find that the absolute negative effect is significantly stronger than the absolute positive effect on experience products.

Consolidating literature about the effects of online customer reviews on online sales of books and similar types of products, hypothesis 1a, 1b and 1c are formulated as follows:

Hypothesis 1a: *The percentage of positivity in online customer reviews has a significant positive effect on Amazon.com's book sales.*

Hypothesis 1b: *The percentage of negativity in online customer reviews has a significant negative effect on Amazon.com's book sales.*

Hypothesis 1c: *The absolute negative effect is significantly stronger than the absolute positive effect on Amazon.com's book sales.*

2.3.2. Effects of the percentage of positive and negative sentiment in expert reviews on book sales

The differential impacts of positive and negative traditional earned media on book sales have also been examined in existing literature. Previously, research only documented the positive effect of positive publicity (e.g. McCracken, 1989; Holbrook, 1999), and the negative effect of negative publicity (e.g. Tybout et al., 1981; Wyat & Badger, 1984). However, a new line of research has emerged and called for attention on the positive effect of negative expert reviews (e.g. Eliashberg & Shugan, 1997; Basuroy et al., 2003). Sorensen and Rasmussen (2004) indicate that "any publicity is good publicity", meaning both positive and negative expert reviews lead to increases of hardcover fiction books' sales. Nevertheless, the authors note that the positive effect of positive expert reviews on the sales of hardcover fiction books is significantly stronger than the positive effect of negative expert reviews.[2] Berger, Sorensen and Rasmussen (2010) build upon the work of Sorensen and Rasmussen (2004) and investigate the differential effects of negative reviews on books with different popularity. According to Berger et al. (2010), *New York Times* expert reviews with negative sentiment negatively affect sales of books by established authors, but have a positive impact

---

[1] Cui et al. (2012) classify online customer reviews with 1 – 2 stars, 3 stars and 4 – 5 stars respectively as having negative, neutral and positive ratings. The researchers measure the valence of online customer reviews using frequencies of numeric ratings to generate the percentages of positive, neutral and negative reviews. Yet, the neutral category is excluded from their analysis.

[2] In Sorensen and Rasmussen (2004), each book with *P* positive sentences, *N* negative sentences and *Z* neutral sentences receives a score of *P/(P+N).* In their econometric analysis, reviews are classified as positive is their score is above 67%, negative if it is less than 33%, and neutral otherwise.

on that of books by less known authors.[3,4] Overall, these studies suggest that the absolute positive effect of positive expert reviews have a stronger effect on book sales than the absolute positive effect of negative expert reviews. Additionally, the less popular the books are, the greater the likelihood of a positive effect of the negative sentiment in expert reviews on them is. These findings lead to the formulation of hypothesis 2a, 2b, 2c, 2d and 3a as follows:

Hypothesis 2a: *The percentage of positivity in expert reviews has a significant positive effect on Amazon.com's book sales.*

Hypothesis 2b: *The percentage of negativity in expert reviews has a significant positive effect on Amazon.com's book sales.*

Hypothesis 2c: *The absolute positive effect of the percentage of positivity in expert reviews is significantly stronger than the absolute positive effect of the percentage of negativity in expert reviews on Amazon.com's book sales.*

Furthermore, the differential effects of negative reviews on popular and less popular books imply an interaction effect:

Hypothesis 2d: *There is a negative interaction effect between the percentage of negativity in expert reviews and book popularity on sales.*

In spite of hypothesis H2b, the studies cited above posit that negative expert reviews may also have a negative effect on sales. Hence, an anti-thesis hypothesis to H2b is formulated:

Hypothesis 3: *There is a negative effect of the percentage of negativity in expert reviews on Amazon.com's book sales.*

2.3.3. Weighing the effect of online customer reviews against that of expert reviews on online book sales

Comparing the effectiveness of traditional and social earned media in influencing product sales, scholars put forward that the latter is a more effective vehicle for boosting sales. Huang and Chen (2006) investigate the herding phenomenon in online product choice and their experiment reveals that the choice made by participants are more positively influenced

---

[3] Berger et al. (2010) use the same calculation of the percentages of sentiments as Sorensen and Rasmussen (2004). However, they categorize reviews as negative if the ratio is below 50% and positive otherwise.

[4] Berger et al. (2010) measure book popularity via a categorical variable (1 = *authors who published one or fewer books prior to the book chosen for analysis*, 2 = *authors who have published between two and nine*, 3 = *authors who have published 10 or more*).

by the recommendation of other consumers than those of an expert. Additionally, Stephen and Galak (2012) propose that for per-event sales, traditional earned media (i.e. newspapers, magazines, television and radio) are more influential than that of social earned media (i.e. blogs and discussion forum posts). Nevertheless, the impact of social earned media on sales has a greater elasticity than traditional earned media. However, none of these studies clarify whether different sentiments expressed in the two types of earned media play a role in the difference between their influences on sales. Also, they both use volume of media content as their measure. Therefore, it is not yet known about the variation between the impacts of positive and negative sentiments expressed in traditional and social earned media. In light of these findings, together with the arguments for prior hypotheses, this research formulates hypothesis 4a and 4b as follows:

Hypothesis 4a: *The effect of the percentage of positivity in online customer reviews is significantly stronger (more positive) than that of the percentage of positivity in expert reviews on Amazon.com's book sales.*

Hypothesis 4b: *The effect of the percentage of negativity in online customer reviews is significantly stronger (more negative) than that of the percentage of negativity in expert reviews on Amazon.com's book sales.*

2.4. Moderating effect of book category

Earned media do not have equal sales effect on different products, and product characteristics play a crucial role in moderating its influence (Zhu & Zhang, 2010). Correspondingly, some researchers show that product type determines customers' search behavior and use of online information sources, which in turn results in sales distribution (e.g. King & Balasubraminian, 1994; Mudambi & Schuff, 2010). As a result, in order to examine the impact of online customer reviews and expert reviews on Amazon.com's book sales, it is important to specify the product type of books and explain how reviewers evaluate such products (Zhu & Zhang, 2010).

2.4.1. Defining book categories

Nelson's (1970) classification of search and experience goods is the product type classification most often used in research in order to understand the effect of presentation format (e.g. Huang, Lurie & Mitra, 2009; Xu, Chen & Santhanam, 2015). Search goods refer to products or services whose features are assessed before purchase and consumption. On the contrary, the characteristics of experience goods are unknown to customers until purchase and use of products (Nelson, 1970). However, Nelson's (1970) classification scheme presents conceptual and practical difficulties (Laband, 1991), for which this study has managed to identify two main causes. First, the Internet has significantly lowered the cost of gathering and sharing information. At the same time, it has opened up new ways for customers to learn about products before purchase and use. As a consequence, the distinction between search and experience goods has been blurred in online environment (Huang et al., 2009). Second, there is a lack of consensus among scholars over the definition of search and experience goods (Laband, 1991). To exemplify, Nelson (1970) classifies food as experience good even though a lot of information about it is available for customers' evaluation before purchases such as nutrition facts, weight, color, shape or even smell (Laband, 1991). This is to show that it is dificult to strictly apply Nelson's (1970) product classification because most of the goods have a mixtured of both search and experience good attributes.

Considering the above mentioned argumentation, it is more logical to regard search versus experience good as the two extreme points of product classification (Laband, 1991). Nelson (1974, 1976, 1981) confirms this criticism and – in his subsequent work – refines his classificatin scheme in which a product is categorized based on the relative amount of its search and experience attributes. This means, search goods are defined as those having mainly product attributes whose full information can be known to customers before purchase. Experience goods are prominently defined by attributes that are difficult to evaluate until purchase and use of products. Compared to the case of search goods, it is costlier and more difficult to perform information search regarding the quality of experience goods (Nelson, 1974, 1976, 1981). The information search for experience goods is characterized by a high degree of dependence on word-of-mouth communication. This search is motivated by the need for information about product experience of others in order to evaluate the products' worth, hence reducing purchase uncertainty (Nelson, 1974, 1976, 1981; Bolton, Katok & Ockenfels, 2004; Forsythe & Shi, 2003; Pavlou & Gefen, 2004).

In alignment with Nelson's (1974, 1976, 1981) refined classification scheme, this study categorizes the selected books (for the analysis) as either books with more search attributes (category 1) or books with more experience good attributes (category 2) (Table 1). In order to determine the book sections of a specific category, the researcher makes use of Lee et al.'s (2011) product objective versus subjective evaluation standards – which bears striking resemblance to Nelson's (1974, 1976, 1981) product classification, with subjective evaluation standards being similar to experience goods attributes, and objective evaluation standards being similar to search good attributes. Moreover, this classification provides supplementary specificities to assess whether the product attributes are search or experience good attributes. Lee el al.'s (2011) work is built upon previous findings which propose that a product is assessed based on its various attributes (Hong & Wyer, 1989; Venkatesh & Agarwal, 2006). Depending on the purpose of purchasing, customers can use either objective or subjective evaluation standards when they examine a product (Lee et al., 2011). According to Lee et al. (2011), the difference between the two standards lies in whether an accepted ranking-based benchmark for evaluation or not is available. For example, in the case of books, price and shipping (and return) options are a few attributes evaluated by objective standards. Cheaper price is commonly considered to better than more expensive one. Free shipping and return is more favorable than having to pay additional costs for delivery service. As there is a generally agreed-upon ranking for the quality of these attributes, they are evaluated by objective standards. In contrast, such attributes as book format, content design, writing style – whose quality is perceived differently due to personal preferences – are evaluated by subjective standards (Lee et al., 2011). To exemplify, some customers may prefer textbook with humorous writing style while others favor academic writing. These subjective evaluations standards resemble experience good attributes as customers can only know about their satisfaction with them after purchase and use. Based on these arguments, in this study, the researcher decided the category of a book section according to whether the majority of its attributes are objective or subjective evaluation standards. In particular, books with more objective evaluation standards will be classified as being of category 1, whereas, those with more subjective evaluation standards category 2.[5]

2.4.2. Moderating effect of book category on the impact of positive sentiments in online customer reviews on book sales

Drawing from the discussion above, books of category 1 have characteristics similar to those of search goods, whereas, books of category 2 resemble experience goods. Therefore, it is possible to make use of previous literature on these two types of goods to discuss the moderating effect of book category in this study. Cui et al. (2012) find that the average ratings of online customer reviews have a greater positive effect on the sales of search goods than on that of experience goods. Similarly, Hao, Ye, Li and Cheng's experiments (2010) indicate that the effect of positive reviews is stronger in the case of search goods than that for experience goods. In contrast, they find no significant difference in the effects of negative reviews on both type of goods.[6] Based on these findings, the research hypothesizes that the effect of positive reviews on the sales of books in category 1 is greater than that of books in category 2, resulting in hypothesis 5 as follows:

[5] Further explanation on categorizing the selected books is provided in the Methods chapter.

[6] Hao et al. (2010) select one-star and five-star reviews from a popular Chinese shopping website respectively as negative and positive reviews.

Hypothesis 5: *The book category moderates the effect of the percentage of positivity in online customer reviews on Amazon.com's book sales. Specifically, the effect of the percentage of positivity in online customer reviews on books of category 1 is significantly stronger than its effect on the sales of books of category 2.*

2.4.3. Moderating effect of book category on the impact of positive sentiment in expert reviews on book sales

There seems to be no prior research directly about or related to the moderation effect of product category on the sales effect of expert reviews. Instead, scholars examine the differential impacts of expert reviews on the sales of different movie genres, which is relatable to the moderating effect of book category on the impact of expert reviews on book sales since both books and movies share a large amount of similarities in product characteristics (Howkins, 2013). To be more precise, drawing from past research insights, the researcher perceives an analogy between books with more search good attributes (category 1) and widely-released movies (and mainstream movies), and between books with more experience good attributes (category 2) and narrowly-released movies (and arthouse movies) (Reinstein & Snyder, 2005; Gemser & Oostrum, 2007).

Reinstein and Snyder (2005) examine the influence effect of film critics' reviews on opening weekend box office revenue.[7] Their findings suggest that film critics' reviews have a significant positive influence effect on the opening weekend box office revenue of narrowly-released movies. Contrarily, they find no significant influence effect for widely-released movies or popular genres such as action movies or comedies. Equivalently, Gemser and Oostrum (2007) examine whether the impact of critical film reviews on mainstream movies differs from that on art house films. The research results infer that film critics' reviews, irrespective to their positive and negative sentiment, positively influence consumers' demand for art house movies. However, critics' film reviews only serve as a predictor for the financial success of mainstream movies; consumers of mainstream movies mostly rely on other sources of information rather than critical film reviews (Gemser & Oostrum, 2007). A possible explanation for these findings is that consumers receive more quality signals for large budget action movies and comedies from press reports and advertising. In contrast, they have less information to evaluate art house movies in advance since most of these films have limited marketing budget (Reisntein & Snyder, 2005).

[7] The authors apply Eliashberg and Shugan's (1997) definition of influence effect which means the causal effect of reviews on demand holding review quality constant.

Given differential effects of expert reviews on a different type of consumable media, the findings inform hypothesis 6:

Hypothesis 6: *The book category moderates the effect of the percentage of positivity in expert reviews on Amazon.com's book sales. In particular, the effect of the percentage of positivity in expert reviews on books of category 2 is significantly stronger than its effect on the sales of books of category 1.*

2.5. Conceptual model

The study presents the conceptual model below in order to summarize the relationships among different variables and present most of the hypotheses that will be tested in the analysis.

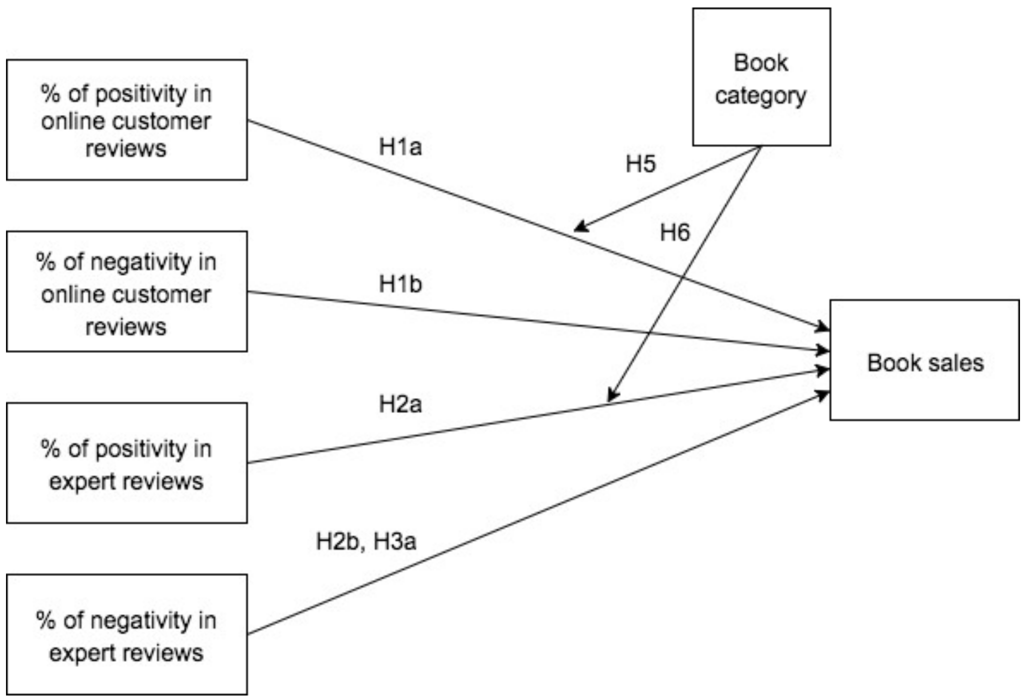


Figure 2. Conceptual model

The remaining hypotheses can be described with the following equations:

H1c: |Negative effect of the percentage of negativity in online customer reviews| > |Positive effect of the percentage of positivity in online customer reviews|

H2c: |Positive effect of the percentage of positivity in expert reviews on book sales| > |Negative effect of the percentage of negativity in expert reviews on book sales|

H3b: Effect of (the percentage of negativity in expert review x book popularity) < 0

H4a: |Effect of <the percentage of positivity> in online customer reviews on book sales| > |Effect of <the percentage of positivity> in expert reviews on book sales|

H4b: |Effect of <the percentage of negativity> in online customer reviews on book sales| >
|Effect of <the percentage of negativity> in expert reviews on book sales|
where | … | indicate absolute value and <…> ∈ {positive, negative}

3. METHOD

3.1. Choice of method

In order to test the hypotheses and hence answer the research question, "*To what extent does traditional (i.e. expert reviews) and social earned media (i.e. online customer reviews) predict Amazon.com's book sales?*", the study took a quantitative research approach. Aliaga and Gunderson (2002) define quantitative research as gathering numerical data which are then analyzed using mathematically based methods (i.e. statistics) in order to explain phenomena. This choice of method was particularly suitable for this study because of three reasons. First, the impact of earned media on product sales had been researched extensively using both qualitative and quantitative methods (e.g. Dellacoras, 2003; Chevalier & Mayzlin, 2006; Zhang et al., 2010; Mudambi & Schuff, 2010; Bao & Chang, 2014), hence providing a firm theoretical basis for the statistical analyses. Second, quantitative research is capable of revealing relationships between variables and explaining their possible causalities (if there is any), which was in line with that the study sets out to test the predictive power of traditional and social earned media, as well as the moderating effect of book category. Third, given the popularity of both expert reviews and online customer reviews on the Internet, there was a vast amount of both textual and numerical data available that can be retrieved and analyzed for the research purpose. Using a large aggregate of data subsequently allowed generalizable results about the effect of expert reviews and Amazon.com online customer reviews on Amazon.com's book sales.

3.2. Steps of analyses

3.2.1. Sentiment analysis

The research was divided into two steps of analyses. To begin with, the study performed a sentiment analysis on selected expert reviews and online customer reviews. Automated sentiment analysis, also known as opinion mining and/or subjectivity analysis, refers to a computational field of research that extract opinions, sentiments and subjectivity in unstructured text in order to identify whether the emotions expressed toward entities (e.g. products, services, organizations, events, etc.) are positive (favorable) or negative (unfavorable) (Pang & Lee, 2008; Liu, 2012). In this study, the sentiment analysis was conducted on document level: its task was to determine the degree of positivity and negativity in each review about the corresponding book sold on Amazon.com and subsequently classify it as being positive, negative or neutral (Liu, 2012). Since the main aim of the sentiment classification is to provide a quick assessment of the predominant opinion about a book, no details about what people liked or disliked about the book are discovered (Liu, 2012). Particularly, in this study, the researcher made use of a lexicon-based sentiment analysis technique (having some natural language processing mechanisms e.g., processing of verb stems) which matched terms with either positive, negative or neutral sentiment based on a precoded wordlist for polarity (Taboada et al., 2011). This lexicon-based approach has been widely applied in many studies to detect sentiment as it requires little training. Additionally, the approach enables the sentiment analysis of both long and short text data (e.g. Bollen et al., 2010; Bollen et al., 2011; Kim, Gilbert, Edwards & Graeff, 2009; Tusmasjan, Sprenger & Sandner, 2010).

3.2.2. LIWC for sentiment analysis

To extract sentiments from expert and online customer reviews, the researcher utilized the text analysis software LIWC 2015 (Linguistics Inquiry and Word Count). The software defines emotional, cognitive and structural components in a text sample based on an internally built dictionary (Pennebaker, Boyd, Jordan & Blackburn, 2015). It relies on simple word count to measure the relative frequency of words related to predefined categories in the dictionary. In other words, it assesses the degree to which positive or negative emotions are present in the text file (Pennebaker et al., 2015). The LIWC dictionary contains a scrupulous amount of 6,400 words, word stems and select emoticons that is constructed by numerous experts with near perfect agreement (Pennebacker et al., 2015). Therefore, the software has been trusted and used extensively in psychology and linguistics to investigate people's thought process, emotions and motivations (Tausczik & Pennebaker, 2010). Specifically, many researchers have employed LIWC to classify positive and negative sentiment in online text data and subsequently examine the predictive power of sentiments expressed on the Web. For example, Tusmasjan et al. (2010) use LIWC to measure political sentiment on Twitter and examine how political sentiment expressed in tweets can predict elections. Krauss et al. (2008) apply it to analyze the degree of positivity in movie discussion within online forums in order to investigate how online sentiment predicts movie success and academy awards. Given its proven credibility and credentials in the field of sentiment analysis, LIWC was a quick and effective digital tool to analyze large text files in this study.

3.2.3. Analysis on sales

After performing the sentiment analysis, the researcher conducted regression analyses, using the statistical analysis software SPSS. In this study, the regression tests were performed on both an aggregated dataset (i.e. group-level) and a disaggregated dataset (i.e. individual-level) because data for review sentiments (i.e. the percentage of positivity and the percentage of negativity in reviews) could be organized at more than one level (Hox, 1995). Specifically, the percentage of positivity and negativity could be measured per book (i.e.

group-level) or per review (i.e. individual-level). There were three advantages of using both datasets which were specific to the case of this study. First, it enabled the researcher to handle group data values for review sentiments (Buxton, 2008). Second, it was useful in overcoming the issue of power failure caused by a small book sample size (Buxton, 2008). Third, it could be used to analyze repeated book sales measures data (Buxton, 2008).

Nevertheless, when testing the hypotheses at both levels, the researcher was aware of the statistical limitations of both aggregated and disaggregated dataset. In the aggregated dataset, different sentiment data values from various subunits are collapsed into fewer sentiment values for fewer high-level sentiment units. Consequently, information is missing, and the power of statistical tests strongly decreases (Hox, 1995). On the other hand, disaggregating data results in some data values of a few high-level units are duplicated for a significantly large number of subunits. Thus, the power of the statistical analysis becomes artificially high (Hox, 1995).

3.3. Sampling

The population of this study includes all books currently available on the online retail website Amazom.com. This online retail website is a rich source of online customer reviews and information about product sales in e-commerce market (Cui et al., 2012). It provides multiple data points regarding customer reviews and book sales (i.e. published date of reviews, ratings, sales ranking) that are essential to the analysis of the study. That these information is well designed and displayed on the website allows, with some effort, quick data retrieval for the research. More importantly, it has one of the largest review community on the Web, providing a rich pool of online customer reviews for sampling.

3.3.1. Scraping book titles, book formats, sales ranking and number of online customer reviews from Amazon.com

A sample of 200 book titles – as well as their book formats, sales ranking and total number of online customer reviews – was drawn from both the *Amazon* bestseller list and the *New York Times* bestseller list. The two rankings were both displayed on Amazon.com, hence being highly visible to a considerable traffic of website visitors. The *Amazon* bestseller list presents the most popular book titles based on their online sales – made via the Amazon.com website – in the previous 24 hours, and the book rankings are updated hourly (www.amazon.com; Chevalier & Goolsbee, 2003). In contrast, the *New York Times* does not provide any explanation on how its bestseller list is generated (Grahl, 2016). Nevertheless, it is widely perceived as a valuable fact-based source for evaluating books (Bao & Chang, 2014). Sorensen (2007) has found evidence that being featured on the *New York Times* bestseller list increases a book's sales. As bestseller books tend to receive vast media attention, the researcher expected that books featured on the *Amazon* bestseller list and the *New York Times* bestseller list to be more likely to receive expert reviews and online customer reviews on Amazon.com than those not, making it easier to scrape an ample amount of reviews for the analysis. Particularly, the researcher first scraped 100 book titles from the Amazon bestseller list and then another 100 from the *New York Times* bestseller list.[8]

Given that the sales ranking of the *Amazon* bestseller list is refreshed hourly, seamless web data scraping within an hour or less was required in order to retrieve correct sales ranking data point. This in turn asked the researcher to already have a complete list of 200 book titles prior to scraping sales ranking data. Because the *Amazon* bestseller list is updated constantly, the researcher purposefully scraped the *New York Times* bestseller list first – which is updated less frequently (i.e. on weekly basis) – within the week of 24th April 2017 but before the date of 28th April 2017, and then moved on to scraping the book titles on the *Amazon* bestseller list – as well as their book formats, sales ranking and total number of online customer reviews – on 28th April 2017 between 1PM and 2PM. Also, during this time period, the researcher manually recorded the Amazon.com's sales ranking and total number of online customer reviews of the 100 book titles that appeared on the *New York Times* bestseller list that week because the list only indicated the book titles and their formats (i.e. hardcover or paperback).

In order to scrape the two bestselling lists, the researcher wrote a Python script, using Python v2.7.13, Python Requests v.2.14.2 (the Non-GMO HTTP library for Python) and BeautifulSoup4 (a Python library for pulling data out of HTML and XML files). The Python script used to scrape the Amazon bestseller list consisted of commands that retrieved data on book titles, book formats (i.e. hardcover, paperback or board book), sales ranking, and total number of online customer reviews. Likewise, the Python script used to scrape the New York Times bestseller list was composed of commands that crawled data on book titles and book formats. The outputs were then parsed and saved as two separate TXT files.

The *Amazon* bestseller list consisted of 16 book titles with board book format. By manual checking, the researcher learned that these board books were children illustration books which neither received expert reviews nor sufficient online customer reviews on Amazon.com. As scraping reviews for these books would little value to the analysis as a whole, they were eliminated from the book sample. Besides, there were 35 book titles that

---

[8] The Amazon bestseller list and the New York Times best seller list each feature 100 book titles.

appear on both bestseller lists whose overlaps were omitted. The deletion resulted in a book sample of N_B = 149 for which expert reviews and online customer reviews were scraped.

3.3.2. Scraping expert reviews and online customer reviews

For each book, 30 expert reviews were obtained through the electronic database of journalistic documents LexisNexis Academic (academic.lexisnexis.nl) throughout the week of 1st May 2017. Since that the sales ranking data were recorded on the 28th April 2017, in order to examine the effect of expert reviews on online book sales, only expert reviews published prior to 27th April 2017 were selected. The same was applied for scraping online customer reviews. LexisNexis Academic's search engine allowed the researcher to enter separate queries for each book, which then returned all relevant search results in English (by selecting '*All English news'* source type). Most of the queries used to find the expert reviews were composed of the book titles placed within double quotation marks (e.g. "Lean In: Women, Work, and the Will to Lead") so that the search engine was signaled to only return articles that contain the exact book titles and avoid extraneous results that may include only one or a few words of the search string. Additionally, the researcher tested out several search strings for each book title in order to determine which keyword combination worked best in returning expert reviews. For books whose titles were likely to overlap with other topics/issues, the researcher entered both the book titles placed in double quotation marks and the authors' names (e.g. "The Most Beautiful: My Life with Prince" Mayte Garcia). Occasionally, the researcher needed to add words such as *novel, biography, non-fiction,* etc. into the search string in order to retrieve the most relevant documents for the study. While scraping expert reviews, the researcher took memos of her own observations about the (non-)selected and (un)available expert reviews which later provided supplementary arguments for the main findings (Appendix I).

Browsing through the search results for a first few queries, the researcher learned that not all search results were qualified as expert reviews. News articles about big-screen/TV series adaptation of some books and interview with the authors' career (or personal life) were brought back alongside the expert reviews. As a result, the researcher had to manually check search results for each book and select documents that fitted the definition of expert reviews in this study. Occasionally, manually scanning through the textual content of the documents was needed to confirm their qualification as expert reviews as it was difficult to determine from only looking at the article titles.

Only 19 out of 149 books have 30 or more expert reviews. Since other books did not meet the minimum requirement of having at 30 expert reviews, their reviews were not retrieved for the analysis. Expert reviews for each of the 19 book titles were downloaded and saved as a TXT file. In total, the amount of scraped expert reviews in this study was 570.

By now, since the specific books that had at least 30 expert reviews were already known, it was not necessary to scrape customer reviews for the entire book sample. Rather, the researcher only scraped online customer reviews from the Amazon.com website for the 19 books whose expert reviews had previously been collected. For each book, 30 online customer reviews on Amazon.com were scraped in the following the week of 8th May 2017. Similar to scraping book titles, the researcher wrote a Python script, using Python v2.7.13, Python Requests v.2.14.2 and BeautifulSoup4.

How customers scroll through customer reviews on the website was taken in account. Amazon.com displays customer reviews in descending publish date order (newest first). Accordingly, the researcher presumed that customers would read the most recent reviews first and then scroll down to the older ones vertically. Therefore, she decided to scrape reviews starting from page one until the quantity of 30 customer reviews published on 27th April 2017 or earlier was met. Prior to scraping, the researcher made use of the website's filter function to display only the customer reviews for the right book formats of the scraped book titles. For example, the book *The Hate U Give* on the *New York Times* bestseller list was of hardcover format. Correspondingly, the researcher chose to only view customer reviews for hardcover format on the website before beginning to scrape. Online customer reviews for each of the book titles were parsed and saved as separate XLSX files. Then, the researcher manually checked and eliminated reviews published before or after the pre-determined duration. In total, there were 570 online customer reviews scraped for the analysis of this study.

3.3.3. Recording sales ranking at (t – 1) time point

Since the research recruited books irrespective to their release data, in order to test hypothesis 2d (*There is a negative interaction effect between the percentage of negativity in expert reviews and book popularity on Amazon.com's book sales.*), sales ranking of the selected book titles at a time point (t – 1) prior time point t = 28th April 2017 needed to be collected. The sales ranking information on 18th April 2017 was available on www.camelcamelcamel.com and was manually recorded using URLs to the chosen books' Amazon.com product pages.

3.4. Operationalization

Book category. In accordance to the theory about product classification (discussed in the Theoretical Framework chapter), the researcher qualitiavely evaluated multiple aspects of each book (e.g. cover design, texual design, book format, book section, etc) in order to determine the category it belongs to. Additionally, a second coder was invited to perform the coding procedure independently in order to increase the reliability of the categorization (Mouter & Noordegraaf, 2012). Comparing the researcher's coding with that of the second coder showed that the second coder was able to reproduce most of the original coding. The coding finally produced the categorization shown in Table 1. The table below presents how the selected books were categorized as either books with more search good attributes (category 1) or books with more experience good attributes (category 2).

Table 1. Classification of 19 books as books with search and experience good attributes

Category 1. Books with search good attributes	**Category 2. Books with experience good attributes**
Shattered: Inside Hilary Clinton's Doomed Campaign	*Option B: Facing Adversity, Building Resilience, and Finding Joy*
How to Win Friends and Influence People	*Zen and the Art of Motorcycle Maintenance*
Sapiens: A Brief History of Humankind	*Hillbilly Elegy: A Memoir of a Family and Culture in Crisis*
Thinking Fast and Slow	*Lean In: Women, Work, and the Will to Lead*
Evicted: Poverty and Profit in the American City	*Women Who Work: Rewriting the Rules for Success*
Originals: How Non-Conformists Move the World	*When Breath Becomes Air*
White Trash: The 400-Year Untold History of Class in America	*1984*
The 48 Laws of Power	*Norse Mythology*
	The Hate U Give
	Prince Charles: The Passions and Paradoxes of an Improbable Life
	Carve the Mark

Book popularity. The independent variable book popularity was measured by the sales ranking values of the selected books which were retrieved from Amazon.com on 18^{th} April 2017. These values were relative indicators of how well the book sales performances had been in the previous 24 hours, hence their real-time popularity on the website. The lower the value of a book's sales rank was, the better its sale performance in the last day was. To illustrate, a book at ranking #1 performed better in terms of sales in the previous 24 hours compared to another at ranking #2.

Book sales. The dependent variable Amazon.com's book sales was measured on a ratio scale. Amazon.com does not release actual book sales, but only their sales ranks. Lee et al. (2011) point out that sales rank cannot be considered as reliable measure for book sales as they only indicate current (or periodical) rather historical sales records. Therefore, in order to measure book sales, this study adopted Lee et al.'s (2011) methodology to transform sales rank of the selected book into a proxy of actual sales. In accordance with Lee et al. (2011), an estimate of book sales (i.e. the number of books sold per book title on Amazon.com) could be calculated on the basis that one percent of sales amount for a book would equal the number of its online customer reviews. In this study, the proxies of actual sales are equal to the total number of online customer reviews published on 27^{th} April 2017 or earlier multiplied by 100.

Sentiment. The independent variables which indicated positive and sentiment (of expert reviews and online customer reviews) were measured on a ratio scale with values expressed in term of percentages. Following Yu, Kaufmann and Diermeier's (2008) methodology, the researcher concatenated all expert reviews and online customer reviews respectively into two separate folder of TXT ad XLSX files to be assessed by LIWC. LIWC assigned to each review a '*posemo'* positive and *'negemo'* negative score – ranging from 0 to 100 – which indicated the percentage of words in the text which were tagged as either positive or negative. The LIWC results then enabled the researcher to construct a variable which measured the percentage of positivity in the reviews and another variable which measured the percentage of negativity in the reviews.

3.5. Validity and reliability

The measurements described above call for a brief discussion about their validity and reliability in the study. Validity is concerned with whether the research measures what it aims to measure. It has three distinct aspects which all need to be warranted in the research, namely: content validity, criterion validity and construct validity (Muijs, 2004). The study met these requirements by incorporating an extensive body of literature to discuss the key concepts in the research question. Similarly, it hypothesized causes and effects based on existing research which had previously produced relevant findings. Given that the sales ranking data on Amazon.com are updated hourly, the research also set up a strict and timely sampling procedure (more details in the Methods chapter) so that accurate data were captured for the analysis. Moreover, by transforming sales ranks into a proxy of actual book sales, the researcher ensured that all statistical results regarding the effect of expert reviews and online customer reviews on Amazon.com's book sales reflected what happened in real life.

The validity of the research was additionally strengthened by choosing LIWC as the software to run sentiment analysis. According to Pennebaker et al. (2015), since the release of the first version of LIWC, its categories and measurements have been proved valid across dozens of psychological domains by hundreds of research. For example, Pennebaker and Francis (1996) test the validity of LIWC scales as part of their experimental study. They ask the participants – first-year undergraduates – to write about their personal thoughts and experiences of attending college. Then, four judges are asked to rate the participants' essays on various emotional dimensions which are designed in accordance with chosen LIWC dictionary scales. Next, using LIWC output and judges' ratings, the researchers test LIWC's external validity by running Pearson's correlational analyses. Their findings support that LIWC measures positive and negative emotions successfully.

Meanwhile, reliability refers to the degree to which statistical results are free of measurement error, which has two primary forms: repeated measurement and internal consistency (Muijs, 2004). In this research, repeated measurement was the relevant form of reliability. One way to achieve the reliability of measurement instruments in this study was by narrowly defining constructs. The researcher extensively elaborated on the definition of expert reviews and online customer reviews so that she was able to stringently select reviews that closely matched the sampling criteria. Data cleaning was also performed so as to further eliminate unsuitable textual content. Another step taken to increase reliability was to have a second coder – who had no prior knowledge of the research goals and objectives – perform qualitative coding on the hypothesized book categories. Resultantly, the second coder was able to independently reproduce the majority of the researcher's coding.

4. RESULTS

4.1. Sample and dataset

An initial book sample of 200 titles was scraped from the *Amazon* bestseller and the *New York Times* bestseller list on 28th April 2017. Out of these 200 titles, there were only 19 books whose online customer reviews and expert reviews were collected for the analysis (N_B = 19). The book sample consisted of 13 hardcover books and six paperback ones.

A review sample of 1,140 reviews was retrieved over a period of two weeks which consists of 570 online customer reviews and 570 expert reviews. The selected online customer reviews and expert reviews had to be published prior to 28th April 2017. After data cleaning, three online customer reviews and seven expert reviews were eliminated from the sample due to their incomplete (or missing) textual content. The deletion resulted in a final review sample of N_R = 1,133 that was used for the analysis, including 567 online customer reviews and 566 expert reviews.

As mentioned in the Methods chapter, the analyses in this study were run on two separate datasets. The first dataset was an aggregated dataset which had 19 research units (i.e. book titles). Each book had one measure of averaged percentage of positivity and negativity respectively in its online customer reviews and expert reviews. Also, it had a single measure of its estimated book sales. In contrast, the second dataset was a disaggregated dataset that treated reviews as research units. This means, there were 1,133 research unit in this dataset. Each review had a single measure of percentage of positivity and negativity in its textual content. Additionally, the estimated sales of its corresponding book title was also recorded; thus, there were repeated measures of book sales in the second dataset.

4.2. Measurements

Review type. In the disaggregated dataset, the variable review type (*'reviewtype'*) was used to separate online customer reviews and expert reviews from each other. It was measured by a binary variable (*1* = online customer reviews, *0* = expert reviews). As already reported above, there are 567 online customer reviews and 566 expert reviews.

Book category. In both datasets, the independent variable was measured by a binary variable (*1* = book with more search good attributes, *2* = book with more experience good attributes). There were eight books of category 1 and 11 books of category 2.

Sentiment. The independent variables positive sentiment and negative sentiment were measured on a ratio scale, and its measures were recorded in the form of percentage.

In the aggregated dataset, each book received four different sentiment scores: the percentage of positivity (*'CR_posem'*) and negativity (*'CR_negem'*) in its online customer reviews, and the percentage of positivity (*'ER_posem'*) and negativity (*'ER_negem'*) in its expert reviews. Each sentiment score was the average of the sentiment scores – positive or negative – for either all online customer reviews or expert reviews a single book had.

Table 2. Range, minimum and maximum values of positive and negative review sentiments in the aggregated dataset

Review	Range	Minimum	Maximum
Online customer reviews			
% of positivity	23.9	5.28	29.18
% of negativity	3.91	.00	3.91
Expert reviews			
% of positivity	3.07	1.42	4.49
% of negativity	2.18	.62	2.8

As shown in Table 2, the percentage of positivity in online customer reviews ranged from 5.28 to 13.19 (*M* = 13.19, *SD* = 5.95), and that of negativity in expert reviews from 0 to 3.91 *(M* = 1.68, *SD* = 1.11). The range of the percentage of positivity in expert reviews was between 1.42 and 4.49 (*M = 2.71*, *SD* = .7), and the percentage of negativity in expert reviews ranged from .62 to 2.8 (*M* = 1.85, *SD* = .62).

In the disaggregated dataset, each review has two sentiment scores: the percentage of positivity (*'posem'*) and negativity (*'negem'*) in its textual content. Their measures were the LIWC outputted scores for positive (*'posemo'*) and negative (*'negemo'*) emotion. The average percentage of positivity in all reviews (*M* = 2.92, *SD* = 2.15) is higher than the average percentage of negativity in all reviews (*M* = 1.97, *SD* = 1.39).

Book popularity. The independent variable book popularity was measured by an ordinal variable (*'rank18'*) whose values equaled the sales ranks of the books on 18th April 2017. In the book sample, the most popular book (i.e. *Shattered: Inside Hilary Clinton's Doomed Campaign*) ranked 1 on Amazon.com on 18th April 2017, whereas, the least popular one (i.e. *Women Who Work: Rewriting the Rules for Success*) ranked 9,167 on the website the same date. On this date, the book sample had ten titles that posited within rank 1 – 100 on Amazon.com and nine ranked lowered than 100.

Book sales. The dependent variable book sales (*'estsales28'*) is an estimate of the actual number of books per book title sold on Amazon.com, measured on a ratio scale. Adopting Lee et al.'s (2011) methodology, the researcher calculated the estimated sales of a book by multiplying its total number of online customer reviews published on 27th April or earlier by 100. In the book sample, the estimated number of total books sold ranged from 3,100 to 165,000. *Women Who Work: Rewriting the Rules for Success* had the least sales of 3,100 books, whereas, *How to Win Friends and Influence People* had the highest sales of 165,300 books. However, the values in the dependent variable *'estsales28'*, in the aggregated dataset, were non-normally distributed, with skewness of .69 (*SE* = .52) and kurtosis of -.92 (*SE* = 1.01). Likewise, those of the equivalent variable, in the disaggregated dataset, were not normally distributed, with skewness of .64 (*SE* = .07) and kurtosis of -.99 (*SE* = .15). This violated the normality assumption that was required for (multiple) regression analyses in this study. For the normality assumption to hold, the researcher used the SQRT() function to transform the dependent variable *'estsales28'* into *'SQRT_estsales28'* in both datasets (See Appendix A1 and A2 for the histograms of variable *'SQRT_estsales28'*). The transformed variables were then used as the dependent variables for measuring estimated number of books sold in all regression analyses.

Interaction term between the percentage of negativity in expert reviews and book popularity. The interaction term between the percentage of negativity in expert reviews and book popularity was measured by a continuous variable. This variable was constructed by multiplying the standardized variables for the percentage of negativity in expert reviews and book popularity. Consequently, the interaction term between the percentage of negativity in expert reviews and book popularity were measured by *'ZER_negemXZrank18'* in the aggregated dataset and *'ZnegemxZrank18'.* In the aggregated dataset, the mean of the interaction term between the percentage of negativity in expert reviews and book popularity was -.45 (*SD* = 1.81). In the disaggregated dataset, the mean of the interaction term between the percentage of negativity in expert reviews and book popularity was -.03 (*SD* = .77).

Interaction term between the percentage of positivity in online customer reviews and book category. The interaction term between the percentage of positivity in online customer reviews and book category was measured by a continuous variable. To construct this variable in the aggregated dataset, the researcher multiplied book category with the standardized variable for the percentage of positivity in online customer reviews. Thus, the interaction term was measured by *'ZCR_posemXbcat'.* The mean of the interaction term between the percentage of positivity in online customer reviews and book category was -.15 (*SD* = 1.17).

In the disaggregated dataset, provided that there was only one sentiment variable (*'posem'*) that measured the percentage of positivity in all reviews, the researcher computed a continuous variable that measured the interaction term between the percentage of positivity in all reviews and book category. By multiplying book category with the standardized variable for the percentage of positivity in the reviews, she created the predicting variable *'ZposemXbcat'.* Then, when using this variable to perform the statistical test, the researcher selected cases where *'reviewtype'* = 1 in order to exclusively examine the moderating effect of book category on the impact that the percentage of positivity in online customer reviews has on Amazon.com's book sales. The mean of the interaction term between the percentage of positivity in reviews and book category was -.03 (*SD* = 1.54).

Interaction term between the percentage of positivity in expert reviews and book category. The interaction term between the percentage of positivity in expert reviews and book category was measured by a continuous variable. For the aggregated dataset, the interaction term was measured by *'ZER_posemXbcat'* which was equal to book category multiplied with the standardized variable for the percentage of positivity in expert reviews. The average value of interaction term between the percentage of positivity in expert reviews and book category was -.08 (*SD* = 1.17).

For the disaggregated dataset, the variable for the interaction term between the percentage of positivity in reviews and book category *'ZposemXbcat'* was used as a predictor in the multiple regression model. The researcher only chose cases where *'reviewtype'* = 0 to specifically study the moderating effect of book category on the impact that the percentage of positivity in expert reviews has on Amazon.com's book sales.

4.3. Hypothesis Testing

4.3.1. Positive effect of the percentage of positivity in online customer reviews on book sales (H1a)

4.3.1.1. Aggregated dataset

To test the effect of the percentage of positivity in online customer reviews on Amazon.com's book sales, a linear regression analysis was performed in which one continuous independent variable (*'CR_posem'*) and one continuous dependent variable

(*‘SQRT_estsales28’*) were used. Prior to running the statistical test, the normality of residuals and homoscedasticity were checked. The normal P-P plot showed that the data appeared to be normally distributed given that it was closely aligned with the diagonal line (Appendix B1). This indicated that the assumption of normality of errors between the independent variable and dependent variable was met. Nevertheless, there were a few data points which deviated significantly from the reference line. The residual scatter plot proved that most scores were posited in the center and scattered along a horizontal line – which satisfied the assumption of homoscedasticity (Appendix B2). The statistical results revealed a positive direction of the effect, $B = 8.87$, but no significant effect of the percentage of positivity in online customer reviews on Amazon.com’s book sales was found, $F(1, 17) = 3.64$, $p = .074$. The test’s insignificance may have been due to too narrow confidence intervals caused by the presence of the outliers. Consequently, hypothesis H1a was rejected.

4.3.1.2. Disaggregated dataset

In the disaggregated dataset, a linear regression test – using only observations from online customer reviews – was also run in order to test hypothesis H1a. In the regression model, the independent variable was the percentage of positivity in online customer reviews (*‘posem’*) and the dependent variable was Amazon.com’s book sales (*‘SQRT_estsales28’*). The required assumptions for normality of errors were met. The scatter plot displayed a (almost) linear relationship between the two variables *‘posem’* and *‘SQRT_estsales28’* (Appendix B3). In addition, the residual scatter plot took approximately the shape of a rectangle, and most scores clustered between the -1 and zero point. This indicated that the homoscedasticity assumption was mostly met (Appendix B4). The results suggested that the linear regression model was significant, $F(1, 565) = 9.56$, $p = .002$. However, it explained marginally 1.7% of the variation in Amazon.com’s book sales, $R^2 = .02$. The percentage of positivity in online customer reviews a significant, positive predictor of Amazon.com’s book sales, $b^* = .13$, $t = 3.09$, $p = .002$, *CI 95%* [.04, .16].[9] There were .83 more books sold on Amazon.com for each additional percent increase in positivity in online customer reviews. Therefore, hypothesis H1a was supported.

---

[9] SPSS only returns the 95% confidence intervals for the unstandardized coefficient *B*. In order to “trick” SPSS into calculating the 95% confidence intervals for the standardized coefficient *b**, the researcher converted the variables *‘SQRT_estsales28’* and *‘posem’* into Z-scores (In SPSS, Analyze > Descriptives Statistics > Descriptives…, select “Save standardized values as variables”). Then, she ran the linear regression model with two variables *‘ZSQRT_estsales28’* and ‘*Zposem’* in order to obtain the reported 95% CIs for *b**.

4.3.2. Negative effect of the percentage of negativity in online customer reviews on book sales (H1b)

4.3.2.1. Aggregated dataset

A linear regression test was run performed to investigate the effect of the percentage of negativity in online customer reviews on Amazon.com's book sales. The regression model used Amazon.com's book sales (*'SQRT_estsales28'*) as dependent variable and the percentage of negativity in online customer reviews ('*CR_negem'*) as independent variable. Both assumptions for normality of errors and constant error variance were partially satisfied. The data were distributed near to the diagonal line but in an S-shaped pattern, providing evidence for large errors in both directions (Appendix C1). The residual scatter plot also showed a random displacement of scores with no clustering and they were distributed in a rectangle pattern (Appendix C2). The results provided no evidence for a significant effect of the percentage of negativity in online customer reviews on Amazon.com's book sales, $F(1, 17) = .03$, $p = .864$. Still, they supported the hypothesized negative direction of the effect, $B = -4.78$. However, its insignificance rendered hypothesis H1b unsupported.

4.3.2.2. Disaggregated dataset

A linear regression test was also performed to test hypothesis H1b in the disaggregated dataset. The analysis had Amazon.com's book sales (*'SQRT_estsales28'*) as dependent variable and the percentage of negativity in online customer reviews (predictor is *'negem'*), and only used observations for online customer reviews (*'reviewtype'* = 1). The normal P-P plot showed an S-shaped pattern of the scores, indicating that the excessive kurtosis of the residuals. Therefore, it was concluded that the assumption for normality of errors was not met (Appendix C3). However, the assumption for homoscedasticity was partially met. Data were concentrated around the zero point in a rectangle pattern though the data points seemed to cluster. Considering that real data rarely has normally distributed errors, it was still possible to perform the linear regression provided that the other assumption was satisfied (Garba, Oyejola, & Yahya, 2013). The results showed that there was no significant effect of the percentage of negativity in online customer reviews on Amazon.com's book sales, $F(1, 565) = .03$, $p = .871$ although the negative direction of the effect was supported, $B = -.16$. As a result, hypothesis H1b was rejected.

4.3.3. Comparing the size of the absolute positive effect of the percentage of positivity and that of the absolute negative effect of the percentage of negativity in online customer reviews on book sales (H1c)

4.3.3.1. Aggregated dataset

Contrary to what had been hypothesized (i.e. $|\beta_2| > |\beta_1|$), the absolute positive effect of the percentage of positivity in online customer reviews appeared to be stronger than the absolute negative effect of the negativity in online customer reviews, $|\beta_1|$ = 8.87, $|\beta_2|$ = 4.78, $|\beta_1| > |\beta_2|$. A z-test for regression coefficients was performed in order to compare the regression coefficient for the effect of the percentage of positivity in online customer reviews on Amazon.com's book sales to that of the percentage of negativity in online customer review on Amazon.com's book sales. The z-score (with mean of 0 and standard deviation of 1) was computed using the following equation:

$$z - score = \frac{|\beta_1| - |\beta_2|}{\sqrt{s_1^2 - s_2^2}}$$

where β_1 and β_2 were the two regression coefficients being compared and s_1 and s_2 were their reported standard errors. The result showed that the difference between the coefficients was not statistically significant, z = .14, p = *.444*, one-tailed. As a consequence, hypothesis H1c was rejected.

4.3.3.2. Disaggregated dataset

Similar to the case of the aggregated dataset, the absolute negative effect of the percentage of negativity in online customer reviews was reportedly weaker than the absolute positive effect of the percentage of positivity in online customer reviews, $|\beta_1|$ = .83, $|\beta_2|$ = .16, $|\beta_1| > |\beta_2|$. However, the z-test results indicated that the difference between the regression coefficients was not statistically significant, z = .02, p = .492, one-tailed. Accordingly, hypothesis H1c was rejected.

4.3.4. Positive effect of the percentage of positivity in expert reviews on book sales (H2a)

4.3.4.1. Aggregated dataset

In the aggregated dataset, the effect of the percentage of positivity in expert reviews on Amazon.com's book sales was examined using a linear regression in which the independent variable was the percentage of positivity in expert reviews (*'ER_posem'*) and the dependent variable was Amazon.com's book sales (*'SQRT_estsales28'*). The assumption of normal distribution of residuals was not satisfied. Although the scores lied near to the reference line, several outliers were present in both directions (Appendix D1). However, the assumption for homoscedasticity was met. The scores were randomly distributed which somewhat formed a rectangle and scattered along a horizontal line (Appendix D2). The results showed the regression model was not significant, $F(1, 17) = 1.49$, $p = .239$. It was possible that the outlier effect undermined the statistical significance. Nevertheless, the positive direction of the effect was – though insignificant – supported, $B = 50.93$. As a result, hypothesis H2a was partly supported.

4.3.4.2. Disaggregated dataset

The hypothesis was tested with a linear regression test in the disaggregated dataset. The regression model used Amazon.com's book sales (*'SQRT_estsales28'*) as dependent variable and the percentage of positivity in expert reviews (predictor is *'posem'*; *'reviewtype'* = 0) as independent variable is significant. Before performing the linear regression, the assumptions for normality of errors and homoscedasticity were checked. The normal P-P plot displayed an obvious S-shaped pattern of score deviations, which indicated that residuals were non-normally distributed (Appendix D3). Nonetheless, the assumption for homoscedasticity was largely met. The data centered around the zero point and across a horizontal line, with some degree of clustering (Appendix D4). The linear regression model was significant, $F(1, 564) = 4.9$, $p = .027$. It helped explain 0.9% of the variance in Amazon.com's book sales, $R^2 = .009$. The percentage of positivity in expert reviews was a significant, positive predictor of Amazon.com's book sales, $b^* = .09$, $t = 2.21$, $p = .027$, *CI 95%* [.07, 1.19].[10] For each additional percent increase in positivity in expert reviews, there were 5.29 more books sold on Amazon.com. Therefore, hypothesis H2a was supported.

4.3.5. Negative or Positive effect of the percentage of negativity in expert reviews on book sales (H2b and H3a)

4.3.5.1. Aggregated dataset

The effect of the percentage of negativity in expert reviews on Amazon.com's book sales was tested using a linear regression model with of Amazon.com's book sales *('SQRT_estsales28')* as dependent variable and the percentage of negativity in expert reviews (*'ER_negem'*) as independent variable. The errors were normally distributed as the normal P-P plot showed that scores fell close to the diagonal reference line (Appendix E1). Additionally, the residual scatter plot revealed a random displacement of scores in a rectangle shape, indicating that the homoscedasticity assumption was met (Appendix E2). On the one hand, the results suggested no significant effect of the percentage of negativity

[10] See footnote 10 for how to calculate of CIs for the standardized coefficient in SPSS.

in expert reviews on Amazon's book sales, *F(*1, 17) = .69, *p* = .419. Furthermore, the unstandardized coefficient indicated that the effect has a negative direction, *B* = -40.09. these findings provided no support for hypothesis H2b and thus it was rejected. On the other hand, despite that the effect was insignificant, its negative direction partially supported for the anti-thesis hypothesis H3.

4.3.5.2. Disaggregated dataset

A linear regression test was also used to test the effect of the percentage of negativity in expert reviews on Amazon.com's book sales in the disaggregated dataset. The regression model had Amazon.com's book sales *('SQRT_estsales28'*) as dependent variable and the percentage of negativity in expert reviews (predictor is *'negem'*; *'reviewtype'* = 0) as independent variable. The points in the normal P-P plot formed a seemingly S-shaped pattern which proved the presence of large outliers in both directions (Appendix E3). Thus, this showed that the errors were not entirely normally distributed and thus the assumption for normality of errors was partially met. Similarly, the homoscedasticity assumption was partially satisfied. The scatter plot showed that all scores were distributed close to the zero point but there was evidence of clustering (Appendix E4). The effect of the percentage of negativity in expert reviews on Amazon.com's book sales was found significant, $F(1, 564)$ = 8.81, p = .003. It accounted for 1.5% of the differences in Amazon.com's book sales, R^2 = .02. The percentage of negativity in expert reviews was a significant, negative predictor of Amazon.com's book sales, b^* = -.12, t = -2.97, p = .003*, CI 95%* [-.58, -.12].[11] For each additional percent increase in negativity in expert reviews, there are 10.91 fewer books sold on Amazon.com. These results led to that hypothesis H2b is rejected. However, they supported the anti-thesis hypothesis H3.

4.3.6. Comparing the size of the absolute positive effect of the percentage of positivity in expert reviews and that of the absolute negative effect of the percentage of negativity in expert reviews on book sales (H2c)

4.3.6.1. Aggregated dataset

Hypothesis H2c posited that the absolute positive effect of the percentage of positivity in expert reviews was significantly stronger than the absolute positive effect of the negativity in expert reviews. However, previous statistical results suggested that the percentage of negativity in expert reviews had a negative effect on Amazon.com's book sales (see section 4.3.5.1). Thus, the researcher was unable to perform any statistical analysis to test the stated hypothesis H2c. In other words, hypothesis H2c was rendered unsupported.

[11] See footnote 10

Nonetheless, a z-test for regression coefficients was still conducted to investigate whether the size of the absolute positive effect of the percentage of positivity in expert reviews was significantly stronger than that of the percentage of negativity in expert reviews. Subsequently, the z-test results showed that the difference between the two regression coefficients was insignificant, $z = 0.17$, $p = .436$, one-tailed.

4.3.6.2. Disaggregated dataset

Comparable to the case of the aggregated dataset, earlier results in this study confirmed the negative effect of the percentage of negativity in expert reviews on Amazon.com's book sales (see section 4.3.5.2), which led to rejecting hypothesis H2b. Given that hypothesis 2c was formulated upon the assumingly significance of hypothesis H2b, it was not possible to run any statistical analysis to test hypothesis H2c. Hence, hypothesis H2c was rejected. Alternatively, the researcher performed a z-test for regression coefficients to compare the absolute positive effect of the percentage in positivity in expert reviews and the absolute negative effect of the percentage in negativity in expert reviews. The results proved that the difference between the two regression coefficients was not statistically significant, $z = -1.28$, $p = .098$, one-tailed.

4.3.7. Negative interaction effect between the percentage of negativity in expert reviews and book popularity on book sales (H2d)

4.3.7.1. Aggregated dataset

The interaction effect between the percentage of negativity in expert reviews and book popularity on Amazon.com's book sales was scrutinized using an OLS regression. In this regression model, the percentage of negativity in expert reviews (*'ZER_negem'*), book popularity (*'Zrank18'*) and the interaction between the percentage of negativity in expert reviews and book popularity (*'ZER_negemXZrank18'*) were the independent variables, and Amazon.com's book sales (*'SQRT_estsales28'*) was the dependent variable. The assumptions for normality of errors and homoscedasticity were examined before running the regression. The normal P-P plot revealed that the points formed an almost linear pattern, proving the approximately normal distribution of the residuals (Appendix F1). However, the residual scatter plot showed that some degree of heteroscedasticity was present. The points were largely concentrated in the center but seemed to be right-skewed distributed (Appendix F2). The results suggested that the OLS regression model was not significant, $F(3, 15) = 3.17$, $p = .055$. The results suggested that there was no significant interaction effect between the percentage of negativity in expert reviews and book popularity on Amazon.com's book sales. Nonetheless, the reported p-value showed that the regression model was considerably close to be statistically significant, which indicated that hypothesis H2d may be supported in a larger sample size.

4.3.7.2. Disaggregated dataset

In the disaggregated dataset, an OLS regression test was also performed to scrutinize the interaction effect between the percentage of negativity in expert reviews and book popularity on Amazon.com's book sales, using only observations for expert reviews (*'reviewtype'* = 0). The percentage of negativity in expert reviews (predictor is *'Znegem'*), book popularity (*'Zrank18'*) and the interaction between the percentage of negativity in expert reviews and book popularity (*'ZnegemXZrank18'*) were the independent variables, and Amazon.com's book sales *('SQRT_estsales28')* was the dependent variable of the regression model. The normal P-P plot showed that the data was distributed in a linear fashion, thus satisfying the assumption for normality of errors (Appendix F3). However, the assumption for homoscedasticity was not satisfied. The residual scatter plot of the variables produced a cone-like shape, which signaled the presence of heteroscedasticity (Appendix F4). The violation of homoscedasticity may lead to incorrect conclusions about the significance of the regression coefficients. The results indicated that the OLS regression was significant, $F(3, 562) = 51.43$, $p < .001$. The model could explain 21.5% of the total variation in Amazon.com's book sales, $R^2 = .22$. The percentage of negativity in expert reviews, $b^* = -.24$, $t = -6.15$, $p < .001$, *95% CI* [-.89, -.46][12] was a significant, negative predictor of Amazon.com's book sales. There were 82.53 fewer books sold on Amazon.com for each additional percent increase in negativity in expert reviews. Book popularity, $b^* = -.53$, $t = -9.39$, $p < .001$, *95% CI* [-.64, -.42] was a significant, negative predictor of Amazon.com's book sales. For each additional unit in book popularity, there were 64.73 fewer books sold on the website. The interaction between the percentage of negativity in expert reviews and book popularity was not a significant predictor of Amazon.com's book sales, $p = .071$. Therefore, regardless of that, the corresponding unstandardized coefficient confirmed the negative direction of the interaction effect, $B = -33.37$, its insignificance resulted in hypothesis H2d being largely unsupported. Table 3 below provides a summary of the regression model for predicting Amazon.com's book sales from the percentage of negativity in expert reviews, book popularity and the interaction between the percentage of negativity in expert reviews and book popularity.

[12] See footnote 10

Table 3. OLS regression model for predicting Amazon.com's book sales from the percentage of negativity in expert reviews, book popularity and the interaction between the percentage of negativity in expert reviews and book popularity (N = 566)

	Amazon.com's book sales			
	B	*SE*	*b**	*t*
Constant	204.75			
The percentage of negativity in expert reviews	-82.53	13.42	-.24 *	-6.15
Book popularity	-64.73	6.88	-.53 *	-9.39
Interaction between the percentage of negativity in expert reviews and book popularity	-33.37	18.43	-.09 ~	-1.81
R^2	.22			
F	51.43			
df	3			

Significance levels: * p < .001, ~ p > .05

4.3.8. Comparing the absolute positive effect of the percentage of positivity in online customer reviews with that of expert reviews on book sales (H4a)

4.3.8.1. Aggregated dataset

Qualitatively comparing the concerned effect sizes revealed that, contrary to the prediction of hypothesis H4a, the absolute positive effect of the percentage of positivity in expert reviews was stronger than that of the percentage of positivity in online customer reviews, $|\beta_1|$ = 8.87, $|\beta_2|$ = 50,94, $|\beta_2| > |\beta_1|$. To assess if they were statistically different, a z-test for regression coefficients was performed. The results did not provide support for any significant difference between the two effect sizes, z = -1, p = .158, one-tailed. Under these circumstances, hypothesis H4a was rejected.

4.3.8.2. Disaggregated dataset

By the same token, in the disaggregated dataset, the absolute positive effect of the percentage of positivity in expert reviews was stronger (more positive) than the absolute positive effect of the percentage of positivity in expert reviews, $|\beta_1|$ = .83, $|\beta_2|$ = 5.29, $|\beta_2| > |\beta_1|$. This resulted in hypothesis H4a being rejected. Nevertheless, a z-test for regression coefficients was still run in order to assess if there was a significant difference between the two effect sizes. The results indicated that there was a significant difference between the two regression coefficients, z = -1.85, p = .032, one-tailed. Particularly, the absolute positive effect of the percentage of positivity in expert reviews was significantly stronger (more positive) than that of the percentage of positivity in online customer reviews on Amazon.com's book sales.

4.3.9. Comparing the absolute negative effect of the percentage of negativity in online customer reviews to that of expert reviews on book sales (H4b)

4.3.9.1. Aggregated dataset

Hypothesis H4b predicted that the absolute negative effective of the percentage of negativity in online customer reviews was significantly stronger (more negative) than that of the percentage of negativity in expert reviews on Amazon.com's book sales. In contrast, the reported absolute negative effect of the percentage of negativity in expert reviews, qualitatively compared, was stronger (more negative) than that of the percentage of negativity in online customer reviews, $|\beta_1|$ = 4.78, $|\beta_2|$ = 40.09, $|\beta_1| < |\beta_2|$. Thus, hypothesis H4b was unsupported. A z-test for regression coefficients, however, was performed to examine whether the difference between the two effect sizes was statistically significant. The results were not statistically significant, z = -.63, p = .264, one-tailed. As a result, there was no significant difference between the two effect sizes.

4.3.9.2. Disaggregated dataset

Similar to the case of the aggregated dataset, previous tests in this study found that the absolute negative effect of the percentage of negativity in expert reviews was stronger than that of the percentage of negativity in online customer reviews, $|\beta_1|$ = .15, $|\beta_2|$ = 10.91, $|\beta_1| < |\beta_2|$. This contradicted hypothesis H4b, which led to rejecting hypothesis H4b. Nonetheless, the researcher conducted a z-test for regression coefficients to examine if there was a significant difference between the two effect sizes. The results suggested that the difference was statistically significant, z = -2.83, p = .002, one-tailed. Hence, the absolute negative effect of the percentage of negativity in expert reviews was significantly stronger (more negative) than that of the percentage of negativity in online customer reviews on Amazon.com's book sales.

4.3.10. Moderating effect of book category on the impact of the percentage of positivity in online customer reviews on book sales (H5)

4.3.10.1. Aggregated dataset

To test the moderating effect of book category on the effect that the percentage of positivity in online customer reviews had on Amazon.com's book sales, an OLS regression test was performed with book category (*'bcat'*), the percentage of positivity in online customer reviews (*'ZCR_posem'*), the interaction between book category and the percentage of positivity in online customer reviews (*'ZCR_posemXbcat'*) as independent variables and Amazon.com's book sales as dependent variable. Both assumptions for normality of errors and constancy of error variance were satisfied. In the normal P-P plot, the scores were aligned close to the diagonal reference line which suggested that there was normal distribution of errors (Appendix G1). Besides, the residual scatterplot showed that there was random displacement of points which took on the shape of a rectangle with no clustering (Appendix G2). Hence, homoscedasticity was present. The results provided no support for a significant moderating effect of book category on the effect of the percentage of positivity in online customer reviews on Amazon.com's book sales, $F(3, 15) = 1.07$, $p = .391$. The results indicated that book category had no significant moderation effect on the impact of the percentage of positivity in online customer reviews on Amazon.com's book sales. Therefore, hypothesis 5 was rejected.

4.3.10.2. Disaggregated dataset

Hypothesis H5 was again tested with an OLS regression in the disaggregated dataset, using only observations for online customer reviews. The regression model used book category (*'bcat'*), the percentage of positivity in online customer reviews ('Zposem') and the interaction between the percentage of positivity in online customer reviews and book category ('ZposemXbcat') as independent variables and Amazon.com's book sales as dependent variable. The assumptions for normality of residuals was not met because the scores displayed in the normal P-P plot were skewed to the right (Appendix G3). However, the assumption for constancy of error variance was partially satisfied as most of the data was scattered in the center within a rectangle form and relatively across a horizontal line (Appendix G4). The results proved that the OLS regression model was significant, $F(3, 563) = 7.14$, $p < .001$. The regression model explained 3.7% of the differences in Amazon.com's book sales. Book category, $b^* = -.09$, $t = -2.05$, $p = .041$, *CI 95%* [-.17, 0], was a significant, negative predictor of Amazon.com's book sales. The percentage of positivity in online customer reviews, $b^* = .38$, $t = 2.9$, $p = .004$, *CI 95%* [.09, .48], was a significant, positive predictor of Amazon.com's book sales. For each additional percent increase in positivity in online customer reviews, there were 34.89 more books sold on Amazon.com. The interaction between the percentage of positivity in online customer reviews and book category, b^* = -.27, t = -2.09, p = .036, *CI 95%* [-.26, 0], was a significant, negative predictor of Amazon.com's book sales. For each additional point in the interaction between the percentage of positivity in online customer reviews and book category, there were 16.24 fewer books sold on Amazon.com. These results indicated that the book category moderated the impact of the percentage of positivity in online customer reviews on Amazon.com's book sales. Furthermore, the values of corresponding unstandardized coefficients showed that the effect of the percentage of positivity in online customer reviews on Amazon.com's book sales was stronger in book category 1 (i.e. books with more search good attributes), B_1 = 18.65, than it was in book category 2 (i.e. books with more experience good attributes), B_2 = 2.41. These findings supported hypothesis H5. See Figure 3 for the graph of the moderation effect and table 3 for the OLS regression model of moderating effect of book category on the effect of the percentage of positivity in online customer reviews on Amazon.com's book sales.

Table 4. OLS regression model for predicting Amazon.com's book sales (N = 567)

	Amazon.com's book sales			
	B	*SE*	*b**	*t*
Constant	236.38			
Book category	-21.87	10.68	-.09 *	-2.05
The percentage of positivity in online customer reviews	34.89	12.03	.38 *	2.9
Interaction between book category and the percentage of positivity in online customer reviews	-16.24	7.75	-.27 *	-2.09
R^2	.04			
F	7.14			
df	3			

Significance level: * p < .05

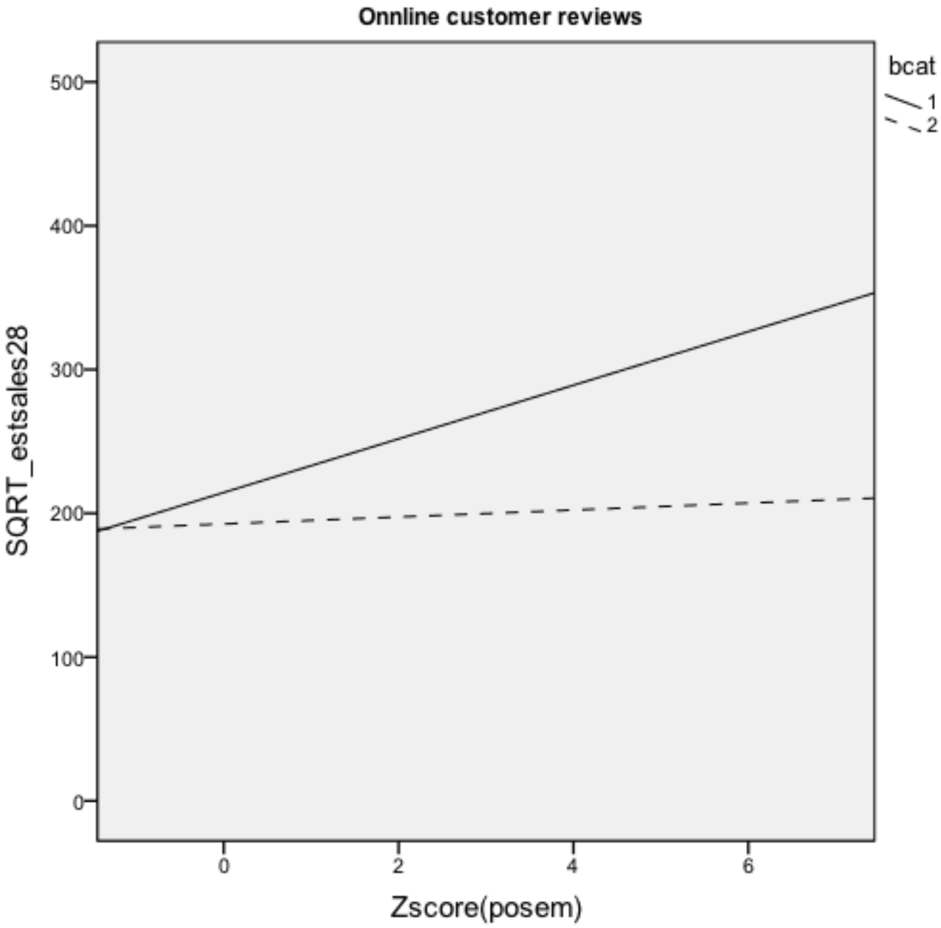


Figure 3. The effect of the positivity in online customer reviews on the sales of books of category 1 and that of category 2

4.3.11. Moderating effect of book category on the effect of the percentage of positivity in expert reviews on book sales (H6)

4.3.11.1. Aggregated dataset

An OLS regression test was run to examine the moderating effect of book category on the effect of the percentage of positivity in expert reviews on Amazon.com's book sales. The regression model used book category (*'bcat'*), the percentage of positivity in expert reviews and the interaction between book category ('*ZER_posem'*) and the percentage of positivity in expert reviews ('*ZER_posemXbcat'*) as independent variables and Amazon.com's book sales (*'SQRT_estsales28'*) as dependent variable. Both assumptions for normality of errors and homoscedasticity were satisfied. The points were close to the diagonal reference line in a linear pattern, which suggested there was normal distribution of residuals (Appendix H1). Meanwhile, the residual scatter plot revealed that the scores were concentrated in the center and distributed in a rectangle pattern. Hence, homoscedasticity was present (Appendix H2). The results provided no support for a significant moderating effect of book category on the effect of the percentage of positivity in expert reviews on Amazon.com's book sales, $F(3, 15) = .75$, $p = .539$. Consequently, hypothesis 6 was rejected.

4.3.11.2. Disaggregated dataset

In the disaggregated, hypothesis H6 was also tested with an OLS regression model of book category (*'bcat'*), the percentage of positivity in expert reviews (predictor is *'Zposem'*; *'reviewtype'* = 0) and the interaction between book category and the percentage of positivity in expert reviews (*'ZposemXbcat'*) as independent variables and Amazon.com's book sales (*'SQRT_estsales28'*) as dependent variable. The assumption for normality of was not satisfied. The normal P-P plot showed that the scores were skewed to the right. Nonetheless, the assumption for homoscedasticity was partially met. The points were scattered around the zero point in the form of a rectangle but distributed along a sloping line. The OLS regression model was found to be significant, $F(3, 562) = 3.77$, $p = .011$. However, none of the predictors in the model are significant: book category, $p = .233$, the percentage of positivity in expert reviews, $p = .315$, the interaction between book category and the percentage of positivity in expert reviews, $p = .71$. A statistically significant regression model but statistically insignificant predictors may have been caused by high degree of multicollinearity. Thus, Variance Inflation Factor (VIP) values were calculated in order to detect multicollinearity. The multicollinearity test suggested that there was no multicollinearity and thus the insignificant effects did not result from high correlations among the predictors. Overall, these statistical findings suggested that the regression model comprised of the main effects and the interaction effect was helpful in predicting Amazon.com's book sales. However, the predictors – put together in the regression model – cancelled each other's predictive power out. Consequently, there was no significant moderating effect of book category on the effect of the percentage of positivity in expert reviews on Amazon.com's book sales. Henceforth, hypothesis H6 was unsupported.

4.4. Summary of statistical results

The table below provides a quick overview of the statistical results produced above.

Table 5. Summary of statistical results

	Hypothesis	Dataset	
		Aggregated	**Disaggregated**
Online customer reviews	H1a	Rejected	Supported
	H1b	Partly supported	Rejected
	H1c	Rejected	Rejected
Expert reviews	H2a	Partly supported	Supported
	H2b	Rejected	Rejected
	H2c	Rejected	Rejected
	H2d	Rejected	Rejected
	H3	Partly supported	Supported
Comparison between online customer reviews and expert reviews	H4a	Rejected	Rejected
	H4b	Rejected	Rejected
Moderating effect	H5	Rejected	Supported
	H6	Rejected	Rejected

5. CONCLUSION

The expansion of online book retail over the past decades has been impressive (Cader, 2001) and it – as an e-commerce sector – will continue to gain ground in the future. In order to keep up with its fast growth, retailers need to a more thorough understanding about the factors that influence online book sales performance. Insofar, previous scholarship has successfully identified social earned media (i.e. online customer reviews) and traditional earned media (i.e. expert reviews) as two significant predictors of online book sales (e.g. Chevalier & Goolsbee, 2003; Chevalier & Mayzlin, 2006; Zhu & Zhang, 2010; Lee et al., 2014). Nevertheless, the fact that these two types of earned media interact with each other while influencing online book sales has not been taken into consideration in past research. Equally important, studies that pay attention to the impact of review sentiments on online sales are scarce (Archak et al., 2011), even though there is evidence that emotions expressed in textual content of reviews can significantly affect customers' purchase decisions (Liu, 2006; Miller et al., 2009; Archak et al., 2011). Acknowledging these shortcomings, the research aimed to offer a broader picture of the interplay between traditional and social earned media in affecting online book sales. Additionally, the study aimed to explore the degree to which book category would moderate the effects of online customer reviews and expert reviews on Amazon.com's book sales. Thus, the research question was formulated as follows: *To what extent do traditional (i.e. expert reviews) and social earned media (i.e. online customer reviews) predict Amazon.com's book sales?* By performing both computational sentiment analysis and hypothesis testing, it produced further insights into the differential effects of online customer reviews and expert reviews on online book sales. Moreover, its findings helped online book retailers tailor their investments on online customer reviews and expert reviews to enhance their online sales.

5.1. Main findings

5.1.1. Effects of online customer reviews and expert reviews on book sales

The research results supported the significance of online customer reviews and expert reviews as predictors of online book sales. To enumerate, the positive sentiment in both online customer reviews and expert reviews had a significant positive effect on Amazon.com's book sales. These findings were aligned with Cui et al. (2012) who argue that positive online customer reviews positively affect online book sales. They also confirmed the findings of prior research on expert reviews which previously proved the positive impact of positive expert reviews (McCracken, 1989; Holbrook, 1999). However, only the negative sentiment in expert reviews had a significant negative effect on Amazon.com's book sales. This result vouched for extant literature which points to the negative impact of negative expert reviews (e.g. Tybout et al., 1981; Wyat & Badger, 1984; Berger et al., 2010). Nevertheless, it did not fully match with the theory of Zhang et al. (2011) and Cui et al. (2012) which assert that negative customer reviews negatively affect online book sales. The insignificant negative effect of the negative sentiment in online customer reviews can be explained by how the reviews are presented on Amazon.com's product page. Amazon.com seemingly displays customer reviews with higher ratings first and places those with lower ratings in subsequent pages. Since the researcher scraped 30 online customer reviews in descending order, it may have led to that online customer reviews with higher percentage of negativity in their textual content were underrepresented in the review sample. Consequently, the predictive power of the percentage of negativity in online customer reviews was downplayed in hypothesis testing. Moreover, in this research, the positive effect of the negative sentiment in expert reviews was not statistically proven, which contradicted the literature (Eliashberg & Shugan, 1997; Basuroy et al., 2003; Sorensen & Rasmussen, 2004). However, the outcome was probable since scholars, in fact, have constantly been producing contradictory results about the relationship between negative sentiment in expert reviews and online book sales.

Besides, there were statistical results that, though insignificant, revealed compelling insights into the differential effects of review sentiments on Amazon.com's book sales. Based on previous research, the expectation was that the negative effect of the negative sentiment in online customer reviews on Amazon.com's book sales would be significantly stronger than the positive effect of the positive sentiment. However, in this research, for online customer reviews, the positive effect of the positive sentiment was found to be stronger than the negative effect of the negative sentiment. This may have been caused by similar reason to that of the insignificant negative effect of the negative sentiment in online customer reviews (discussed above). Contrary to the case of online customer reviews, the positive effect of the positive sentiment in expert reviews was stronger than the negative effect of the negative sentiment. This may have occurred due to the interference of the gatekeepers within the online publishing sector. In detail, before being published, expert reviews already went through editorial which may have been influenced not only by those who work at the newspapers but also book publishers (or distributors) who would like to guard against unfavorable light on their book titles.

Another surprising revelation was that the sentiments in expert reviews had a, though not significant, stronger effect on Amazon.com's book sales than those of online customer reviews. Specifically, the positive effect of the positive sentiment and the negative effect of the negative sentiment in expert reviews on Amazon.com's book sales were stronger than those of online customer reviews. Revisiting the memos taken during the retrieval of expert reviews, the researcher came to realize that many books which received substantial amount of expert reviews were written by acclaimed award-winning authors. According to Dobrescu, Luca and Motta (2013), these authors had already had a considerable base of loyal readers who would prefer to rely on expert reviews for critical evaluation of book content. As a consequence, their book purchase decisions were more likely to be influenced by the sentiments expressed in expert reviews.

Furthermore, in this research, the negative interaction effect between the percentage of negativity in expert reviews and book popularity on Amazon.com's book sales was not statistically significant. This was against the research's expectation that the less popular the books were, the higher the chance of a positive effect of the negative sentiment in expert reviews on them was. The statistical insignificance of the interaction effect may have occurred due to the small amount of the percentage of negativity in the selected expert reviews. That the negative sentiment was not marginally represented in the review sample made it difficult to accurately test hypothesis H2d. In other words, having a more representative sample would allow the researcher to better assess the negative interaction effect between it and book popularity on Amazon.com's book sales.

5.1.2. Moderating effect of book category

The findings in this research proved that book category was a significant moderator of the effect of the positive sentiment in online customer reviews on Amazon.com's book sales. To be more specific, the positive sentiment in online customer reviews had a stronger effect on the online sales of books with more search good attributes than it did on that of books with more experience good attributes. This statistical significance was in a similar vein with past studies by Hao et al. (2011) and Cui et al. (2012) which point out that the impact of the positive sentiment in online customer reviews would be stronger in the case of search goods than it would in the case of experience goods.

On the contrary, the moderating effect of book category on the impact of the percentage of the positive sentiment in expert reviews was statistically insignificant. A possible explanation for this findings was because of the primary sort of book discussion across expert reviews. Particularly, expert reviews single-mindedly concentrated on the technical and artistic features of the books' textual content, as expert reviewers wanted their writings to appear unbiased (though, as mentioned the Theoretical Framework chapter, expert reviews included both factual and personal opinions) (Holbrook, 1999). They minimally elaborated on other aspects of the books which reflected personal preferences and thus helped differentiate between books with more search good attributes and those with more experience good attributes. As a result, the effect of their positive sentiment on the online sales of different book categories did not significantly differ from each other.

5.1.3. Practical implications of main findings

The main findings above revealed the continued relevance of both online customer reviews and expert reviews as influential determinants of online book sales performance. What is more, in spite of customers' reportedly increasing reliance on user-generated content while considering to purchase a book online, the research showed that expert reviews still appeared to play an important role in boosting books' reputation and thus online sales. For this reason, it is advised that online book retailers pay attention to the effects of both types of reviews when promoting their offered book titles. Based on what book category the majority of their books belong to, they need to strategically formulate an actionable financial plan to optimize the combined positive effects of these online communication sources on their sales performance.

5.2. Limitations and recommendations

There were several limitations that may have limited the researcher in arriving at a more generalizable and conclusive answer to the research question, which future research should take into consideration and improve.

To begin with, this research took the initiative to explore the moderating effect of book category on the impact of the percentage of positive sentiment in online customer reviews and expert reviews on Amazon.com's book sales. Since there was no prior research that directly dealt with this topic, the researcher had to draw on general literature on the moderating effect of product category (i.e. search versus experience goods) to formulate the hypotheses, and qualitatively classified the two book categories. Even though the research selectively used renowned scholarship and applied intercoding technique so as to guarantee the validity and reliability of the research findings, flaws could still exist due to its newness. Future research is therefore encouraged to conduct a quantitative pretest to verify the reliability of the qualitative coding results before proceeding with the main data analysis.

Another concern about sampling in this research was caused by the markedly small book sample size, though there was a vast amount of online customer reviews and expert reviews collected for the analysis. This was due to the researcher's limited familiarity with the nature of the publishing industry, which in turn resulted in her not being able to anticipate the scope of difficulty in obtaining expert reviews. Desirably, the research should have had 30 book titles at the minimum to ensure sufficient statistical power during data analysis (Johanson & Brooks, 2010). By doing so, the researcher could have avoided having to use two separate datasets for the analysis, hence being able to generate more consistent and reliable findings. To prevent this shortcoming, future research should consider lowering the required minimum number of online customer reviews and expert reviews to be scraped per book title. Considering e-commerce customers' limited time resource, it is likely that they do not necessarily need to read as many as 30 online customer or expert reviews to reach a (non-)purchase decision.

Still concerning sampling, for convenience and technical reasons, this research only scraped book titles from bestselling lists. Although recruiting book titles helped the researcher produce interesting insights, they were not representative of the wide book selection on Amazon.com. Given the technical feasibility, future research should also include book titles that are not featured on either the *Amazon* and the *New York Times* bestselling list in order to observe the effects of the reviews on the online sales of less known books.

Moreover, this study did not use data of actual sales, but rather a proxy of actual number of books sold per book title as the dependent variable to measure online book sales. The proxies of actual sales were calculated based on Lee et al.'s (2011) assumption that one percent of actual sales of each book would equal the number of its online customer reviews. An apparent disadvantage of this calculation was that not all customers who previous purchased a book would be motivated enough to post an online review. In actual fact, there is evidence that only highly satisfied and extremely disappointed customers are incentivized to write a review for a product bought online (Mudambi & Schuff, 2010). If possible, future research should get in touch with online book retailers and ask for permission to access their real-life sales figures to construct a book sales variable with higher reliability.

REFERENCES

ActivMedia Research. (1999). *Real numbers behind "net profits 1999," sixth annual guide to global e-commerce – executive summary report.* Retrieved from http://www.activmedia.com

Agarwal, J., & Kamakura, W.A. (1995). The economic worth of celebrity endorsers: An event study analysis. *Journal of Marketing, 59*(3), 56–62. DOI: 10.2307/1252119

Aliaga, M., & Gunderson, B. (2002). *Interactive statistics.* Upper Saddle River, NJ: Pearson Education.

Archak, N., Ghose, A., & Ipeirotis, P.G. (2011). Deriving the pricing power of product features by mining customer reviews. *Management Science, 57*(8), 1485–1509. DOI: http://dx.doi.org.eur.idm.oclc.org/10.1287/mnsc.1110.1370

Bao, T., & Chang, T.S. (2014). Why Amazon uses both the New York Times best seller list and customer reviews: An empirical study of multiplier effects on product sales from multiple earned media. *Decision Support Systems, 67*(2014), 1–8. DOI: http://dx.doi.org/10.1016/j.dss.2014.07.004

Bambauer-Sachse, S., & Mangold, S. (2013). Do consumers still believe what is said in online product reviews? A persuasion knowledge approach. *Journal of Retailing and Consumer Services, 20*(4), 373–381. DOI: https://doi.org/10.1016/j.jretconser.2013.03.004

Basuroy, S., Chatterjee, S., & Ravid, S.A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing, 67*(4), 103–117. DOI: 10.1509/jmkg/67.103.18692

Bei, L.T., Chen, E.Y.I, & Widdows, R. (2004). Consumers' online information search behavior and the phenomenon of search vs. experience goods. *Journal of Family and Economic Issues, 25*(4), 449–467. DOI: 10.1007/s10834-004-5490-0

Berger, J., Sorensen, A.T., & Rasmussen, S.J. (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science, 29*(5), 815–827. DOI: http://dx.doi.org/10.1287/mksc.1090.0557

Bollen, J., Mao, H., & Zeng, X.J. (2010). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1–8. DOI: https://doi.org/10.1016/j.jocs.2010.12.007

Bollen, J., Mao, H., & Pepe, A. (2011). *Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena.* Paper presented at the 5th AAAI Conference on Weblogs and Social media. Retrieved June 17, 2017, from www.aaai.org

Bolton, G.E., Katok, E., & Ockenfels, A. (2004). How effective are electronic reputation mechanisms? An experimental investigation. *Management Science, 50*(11), 1587–1602. DOI: http://dx.doi.org/10.1287/mnsc.1030.0199

Brynjolfsson, E., Hu, Y., & Smith, M.D. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online book sellers. *Management Science,* 49(11), 1580–1596. DOI: https://doi.org/10.1287/mnsc.49.11.1580.20580

Buxton, R. (2008). *Statistics: Multilevel modeling.* Retrieved from http://www.statstutor.ac.uk/resources/uploaded/multilevelmodelling.pdf

Chen, P.Y., Dhanasobhon, S., & Smith, M.D. (2008). All reviews are not created equal: The disaggregate impact of reviews and reviewers at Amazon.com. *Social Science Research Network (SSRN).* Advance online publication. DOI: http://dx.doi.org/10.2139/ssrn.918083

Chen, Y., Liu, Y., & Zhang, J. (2012). When do third-party product reviews affect firm value and what can firms do? The case of media critics and professional movie reviews. *Journal of Marketing, 76*(2), 116–134. DOI: http://dx.doi.org/10.1509/jm.09.0034

Chevalier, J., & Goolsbee, A. (2003). Measuring prices and price competition online: Amazon.com and BarnesandNoble.com. *Quantitative Marketing and Economics, 1*(2), 203–222. DOI: 10.1023/A:1024634613982

Chevalier, J.A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345–354. DOI: http://dx.doi.org/10.1509/jmkr.43.3.345

Clay, K., Ramayya, K., & Wolff, E. (2001). Prices and price dispersion on the Web: Evidence from the online book industry. *The Journal of Industrial Economics, 49*(4), 521–539. DOI: 10.1111/1467-6451.00161

Corcoran, S. (2009). Defining earned, owned, and paid media. *Forrester.* Retrieved from http://blogs.forrester.com.

Cui, G., Lui, H.K., & Gui, X. (2012). The effect of online consumer reviews on new product sales. *International Journal of Electronic Commerce, 17*(1), 39–58. DOI: 10.2753/JEC1086-4415170102

Das, S.R., & Chen, M.Y. (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science, 53*(9), 1375–1388. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1070.0704

Dellacoras, C. (2003). The digitalization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science, 49*(10), 1407–1424. DOI: http://dx.doi.org/10.1287/mnsc.49.10.1407.17308

Dobrescu, L.I., Luca, M., & Motta, A. (2013). What makes a critic tick? Connected authors and the determinants of book reviews. *Journal of Economic Behavior & Organization, 96*, 85–103. DOI: https://doi.org/10.1016/j.jebo.2013.08.008

Duan, W., Gu, B., & Whinston, A.B. (2008). Do online reviews matter? – An empirical investigation of panel data. *Decision Support Systems, 45*(4), 1007–2016. DOI: https://doi.org/10.1016/j.dss.2008.04.001

Elberse, A. (2007). The power of stars: Do star actors drive the success of movie? *Journal of Marketing, 71*(4), 102–120. DOI: http://dx.doi.org/10.1509/jmkg.71.4.102

Eliashberg, J., & Shugan, S.M. (1997). Film critics: influencers or predictors. *Journal of Marketing, 61*(2), 68–78. DOI: 10.2307/1251831

Fiske, S.T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology, 38*(6), 889–906. DOI: http://dx.doi.org/10.1037/0022-3514.38.6.889

Forsythe, S.M., & Shi, B. (2003). Consumer patronage and risk perceptions in Internet shopping. *Journal of Business Research, 56*(11), 867–875. DOI: https://doi.org/10.1016/S0148-2963(01)00273-9

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research, 21*(1), 1–31. DOI: https://doi.org/10.1086/209380

Garba, M.K., Oyejola, B.A., & Yahya, W.B. (2013). Investigations of certain estimators for modelling panel data under violations of some basic assumptions. *Mathematical Theory and Modeling, 3*. Retrieved from www.iiste.org

Ganiz, M., Pottenger, W.M., & Yang, X. (2007). *Link analysis of higher-order paths in supervised learning datasets.* Paper presented at the 5th Workshop on Link Analysis, Counterterrorism and Security, SIAM International Data Mining Conference. Minnesota, US.

Gemser, G., Oostrum, M.V., & Leenders, M.A.A.M. (2007). The impact of film reviews on the box office performance of art house versus mainstream motion picture. *Journal of Cultural Economics,* 31(1), 43–63. DOI: 10.1007/s10824-006-9025-4.

Goodall, D. (2009). Owned, bought, and earned media. *Daniel Goodall.* Retrieved from http://danielgoodall.com

Gothberg, J. (2012, April 12). *June Gothberg on using Linguistic Inquiry and Word Count (LIWC) Software.* Retrieved from http://aea365.org/blog/june-gothberg-on-using-linguistic-inquiry-and-word-count-liwc-software/

Grahl, T. (2016, April 2). The truth about the New York Times and Wall Street Journal bestseller lists. *Observer.* Retrieved from http://observer.com/2016/02/the-truth-about-the-new-york-times-and-wall-street-journal-bestseller-lists/

Hao, Y., Ye, Q., Li, Y., & Cheng, Z. (2010). *How does the valence of online consumer reviews matter in consumer decision making? Differences between search goods and experience goods.* Paper presented at the 43rd Hawaii International Conference on System Sciences. Retrieved from http://www.jcomputers.us/vol6/jcp0603-14.pdf

Holbrook, M.B. (1999). *Consumer value: A framework for analysis and research.* Abingdon, UK: Routledge.

Hong, S.T., & Wyer, R.S. (1989). Effects of country-of-origin and product attribute information on product evaluation: An information processing perspective. *Journal of Consumer Research, 16*(2), 175–187. DOI: https://doi.org/10.1086/209206

Howkins, J. (2013). *The creative economy: How people make money from ideas.* London, UK: Penguin.

Hox, J.J. (1995). *Applied multilevel analysis.* Amsterdam, NL: TT-publikaties.

Hu, N., Hok, N.S., & Reddy, S.K. (2013). Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decision Support Systems, 57*(2014), 42–53. DOI: http://dx.doi.org/10.1016/j.dss.2013.07.009

Huang, J.H., & Chen, Y.F. (2006). Herding in online product choice. *Psychology & Marketing, 23*(5), 413–428. DOI: 10.1002/mar.20119

Huang, P., Lurie, N.H., & Mitra, S. (2009). Searching for experience on the Web: An empirical examination of consumer behavior for search and experience goods. *Journal of Marketing, 73*(2), 55–69. DOI: https://doi.org/10.1509/jmkg.73.2.55

Johanson, G.A., & Brooks, G.P. (2010). Initial scale development: Sample size for pilot studies. *Educational and Psychological Measurement, 70*(3), 394–400. DOI: 10.1177/0013164409355692

Kilian, W., Picot, A., Neuburger, R., Niggl, J., Scholtes, K.L., & Seiler, W. (1994). *Aus ökonomischer und juristischer Sicht* [Electronic Data Interchange (EDI)]. Baden-Baden, Germany: Nomos.

Kim, E., Gilbert, S., Edwards, M.J., Graeff, E. (2009). Detecting sadness in 140 characters: Sentiment analysis of mourning Michael Jackson on Twitter. *Web Ecology, 3.* Retrieved from http://www.webecologyproject.org/wp-content/uploads/2009/08/Detecting_Sadness_in_140_Characters2.pdf

King, M.F., & Balasubramanian, S.K. (1994). The effects of experience, end goal, and product type on adoption of preference formation strategy. *Journal of the Academy of Marketing Science, 22*(2), 146–159. DOI: https://doi.org/10.1177/0092070394222004

Krauss, J., Nann, S., Simon, D., Gloor, P.A., & Fischbach, K. (2008). *Predicting movie success and academy awards through sentiment and social network analysis.* Paper presented at the 16th European Conference on Information Systems. Retrieved June 17, 2017, from www.researchgate.net

Kusumasondjaja, S., Shanka, T., & Marchegiani, C. (2012). Credibility of online reviews and initial trust: The roles of reviewer's identity and review valence. *Journal of Vacation Marketing, 18*(3), 185–195. DOI: 10.1177/1356776712449365

Laband, D.N. (1991). An objective measure of search versus experience goods. *Economic Inquiry, 29*(3), 497–509. DOI: 10.1111/j.1465-7295.1991.tb00842.x

Laudon, K.C., & Traver, C.G. (2008). *E-commerce: Business, technology, society.* New York, NY: Pearson.

Lee, J., Lee, J.N., & Shin, H. (2011). The long tail or the short tail: The category-specific impact of eWOM on sales distribution. *Decision Support Systems, 51*(3), 466–479. DOI: http://dx.doi.org/10.1016/j.dss.2011.02.011

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing, 70*(3), 74–89. DOI: http://dx.doi.org/10.1509/jmkg.70.3.74

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1–167. DOI: 10.1.1.244.9480

McCraken, G. (1989). Who is the celebrity endorser? Cultural foundations of the endorsement process. *Journal of Consumer Research, 16*(3), 310–321. DOI: https://doi.org/10.1086/209217

Miller, K.D., Fabian, F., & Lin. S.J. (2008). Strategies for online communities. *Strategic Management Journal, 30*, 305–322. DOI: doi:10.1002/smj.735

Mosendz, P. (2014, May 30). Amazon has basically no competition among online booksellers. *The Atlantic.* Retrieved from https://www.theatlantic.com/business/archive/2014/05/amazon-has-basically-no-competition-among-online-booksellers/371917/

Mouter, N., & Noordegraaf, D.V. (2012). Intercoder reliability for qualitative research. You win some, but do you lose some as well? Retrieved from TUDelf database: http://rstrail.nl/new/wp-content/uploads/2015/02/mouter_-_vonk_noordegraaf_2012.pdf

Mudambi, S.M., & Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon.com. *MIS Quarterly, 34*(1), 185–200. Retrieved from https://ssrn.com/abstract=2175066

Muijs, D. (2004). *Doing quantitative research in education.* Thousand Oaks, London: Sage.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy, 78.* Retrieved from http://www.jstor.org/stable/1830691

Nelson, P. (1974). Advertising as information. *Journal of Political Economy, 82*(4), 729–754. DOI: 10.1086/260231

Nelson, P. (1976). Economic values of advertising. In Y. Brozen (Ed.), *Advertising and Society* (pp. 109–141). New York, NY: New York University Press.

Nelson, P. (1981). Consumer information and advertising. In M. Galatin, & R.D. Leiter (Eds.), *Economics of Information* (pp. 42–77). Boston, MA: N. Nijhoff Publishers.

Nielsen (2016). *2015 U.S. book industry year-end review.* Retrieved from http://www.nielsen.com/us/en/insights/reports/2016/2015-us-book-industry-year-end-review.html

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1–135. DOI: http://dx.doi.org/10.1561/1500000001

Patak, N., Mane, S., Srivastava, J., & Contractor, N. (2007). *Knowledge perception analysis in a social network.* Paper presented at the 5th Workshop on Link Analysis, Counterterrorism and Security, SIAM International Data Mining Conference. Minnesota, US.

Pavlou, P.A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Information Systems Research, 15*(1), 37–59. DOI: http://pubsonline.informs.org/doi/abs/10.1287/isre.1040.0015

Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015.* Austin, TX: University of Texas at Austin.

Pennebaker, J.W., & Francis, M.E. (2010). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion, 10*(6), 601–626. DOI: http://dx.doi.org/10.1080/026999396380079

Reinstein, D.A., & Snyder, C.M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of motive critics. *The Journal of Industrial Economics, 53(1),* 27–51. DOI: 10.1111/j.0022-1821.2005.00244.x

Shrum, L.J., Liu, M., Nespoli, M., & Lowrey, T.M. (2012). Persuasion in the marketplace: How theories of persuasion apply to marketing and advertising. In J.P. Dillard and L. Shen (Ed.), *The SAGE handbook of persuasion: Developments in theory and practice* (pp. 314–330). Thousand Oaks, CA: SAGE.

Skowronski, J.J., & Carlston, D.E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *American Psychological Association, 105*(1), 131–142. DOI: http://dx.doi.org/10.1037/0033-2909.105.1.131

Sorensen, A.T., & Rasmussen, S.J. (2004). Is any publicity good publicity? A note on the impact of book reviews. *NBER Working paper, Stanford University.* Retrieved from http://www.ssc.wisc.edu/~sorensen/papers/bookreviews.pdf

Stephen, A.T., & Galak, J. (2012). The effects of traditional and social earned media on sales: A study of a microlending marketplace. *Journal of Marketing Research, 49*(5), 624–639. DOI: http://dx.doi.org/10.1509/jmr.09.0401

Sundaram, D.S., Mitra, K., & Webster, C. (1998). Word-of-mouth communications: a motivational analysis. *Association for Consumer Research, 25,* 527–531. Retrieved from http://acrwebsite.org/volumes/8208/volumes/v25/NA-25

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics, 37*(2), 267–307. DOI: 10.1162/COLI_a_00049

Tausczik, Y., Faasse, K., Pennebaker, J. W., & Petrie, K. J. (2012). Public anxiety and information seeking following the H1N1 outbreak: blogs, newspaper articles, and Wikipedia visits. *Health Communication, 27*(2), 179–185. DOI: http://dx.doi.org/10.1080/10410236.2011.571759

Trusov, M., Bucklin, R.E., & Pauwels, K. (2009). Effects of word-of-mouth versus traditional marketing: Findings from an Internet social networking site. *Journal of Marketing, 73*(5), 90–102. DOI: http://dx.doi.org/10.1509/jmkg.73.5.90

Turney, P.D. (2002). *Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews*. Paper presented at the 40th Annual Meeting on Association for Computational Linguistics. DOI: 10.3115/1073083.1073153

Tumasjan, A., Sprenger, T.O., & Sandner, P.G. (2010). *Predicting elections with Twitter: What 140 characters reveal about political sentiment.* Paper presented at the 4th International AAAI Conference on Weblogs and Social media. Retrieved May 16, 2017, from www.aaai.org

Tybout, A.M., Calder, B.J., & Sternhal, B. (1981). Using information-processing theory to design marketing strategies. *Journal of Marketing Research, 18*(1), 73 – 79. DOI: 10.2307/3151315

Venkatesh, V., & Agarwal, R. (2003). Turning visitors into customers: A usability-centric perspective on purchase behavior in electronic channels. *Management Science, 52*(3), 367–382. DOI: https://doi.org/10.1287/mnsc.1050.0442

Wigand, R.T. (2006). Electronic commerce: Definition, theory, and context. *The Information Society, 13*(1), 1–16. DOI: http://dx.doi.org/10.1080/019722497129241

Wyat, R.O., & Badger, D.P. (1984). How reviews affect interest in and evaluation of films. *Journalism & Mass Communication Quarterly, 61*(4), 874–878. DOI: 10.1177/107769908406100421

Xu, P., Chen, L., & Santhanam, R. (2015). Will video be the next generation of e-commerce product reviews? Presentation formation and the role of product type. *Decision Support Systems, 73,* 85 – 96. DOI: https://doi.org/10.1016/j.dss.2015.03.001

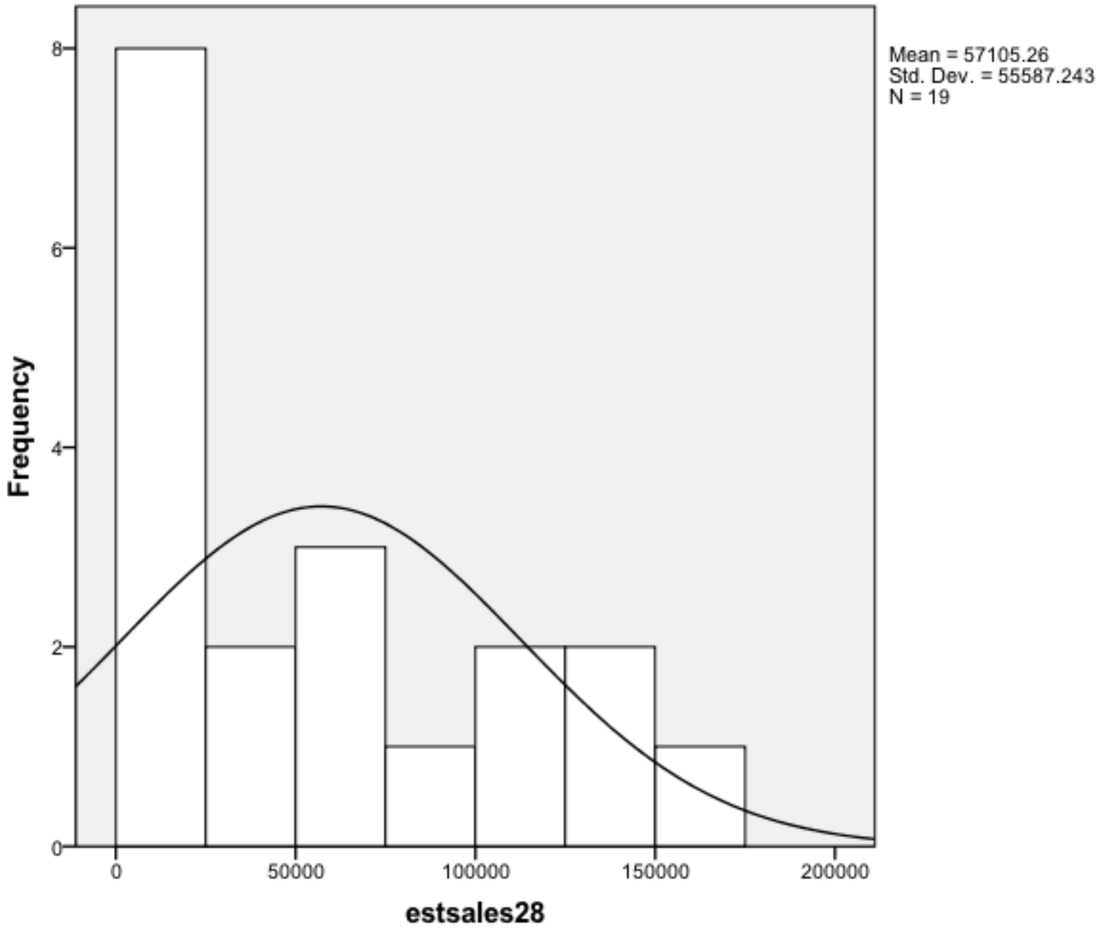
Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems, 55*(4), 919–926. DOI: http://dx.doi.org/10.1016/j.dss.2012.12.028

Yu, B., Kaufmann, S., & Diermeier, D. (2008). Classifying party affiliation from political speech. *Journal of Information Technology & Politics, 5*(1), 33–48. DOI: http://dx.doi.org/10.1080/19331680802149608

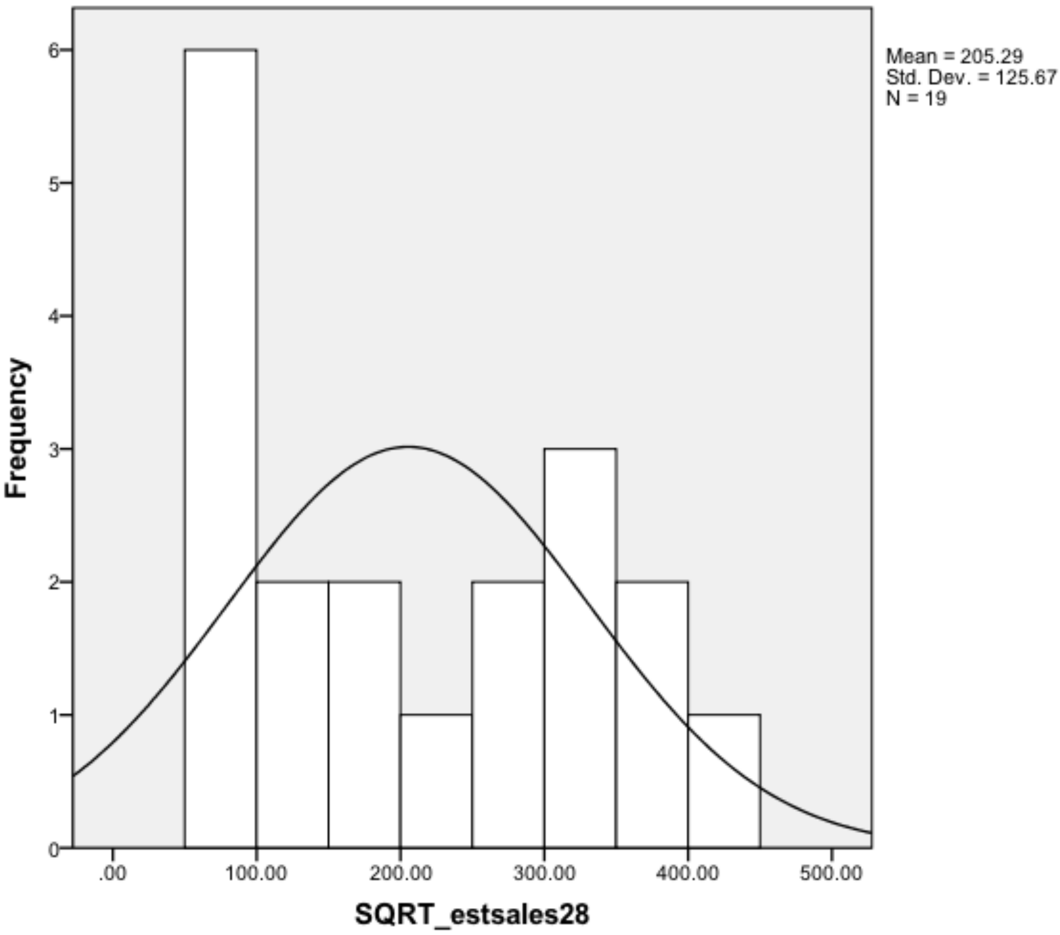
Zhang, J.Q., Cranciun, G., & Shin, D. (2010). When does electronic word-of-mouth matter? A study of consumer product reviews. *Journal of Business Research, 63*(12), 1336–1341. DOI: http://dx.doi.org/10.1016/j.jbusres.2009.12.011

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing, 74*(2), 133–148. DOI: http://dx.doi.org/10.1509/jmkg.74.2.133

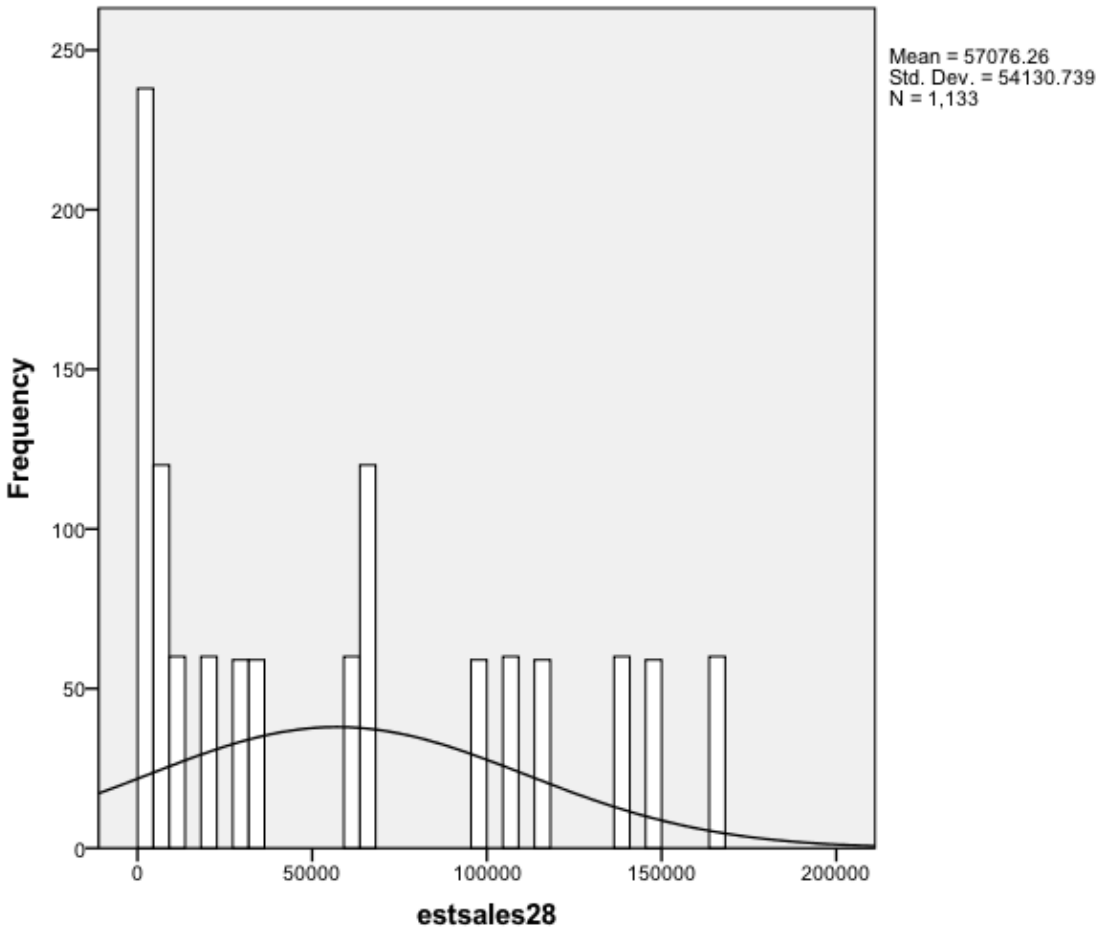
APPENDIX A



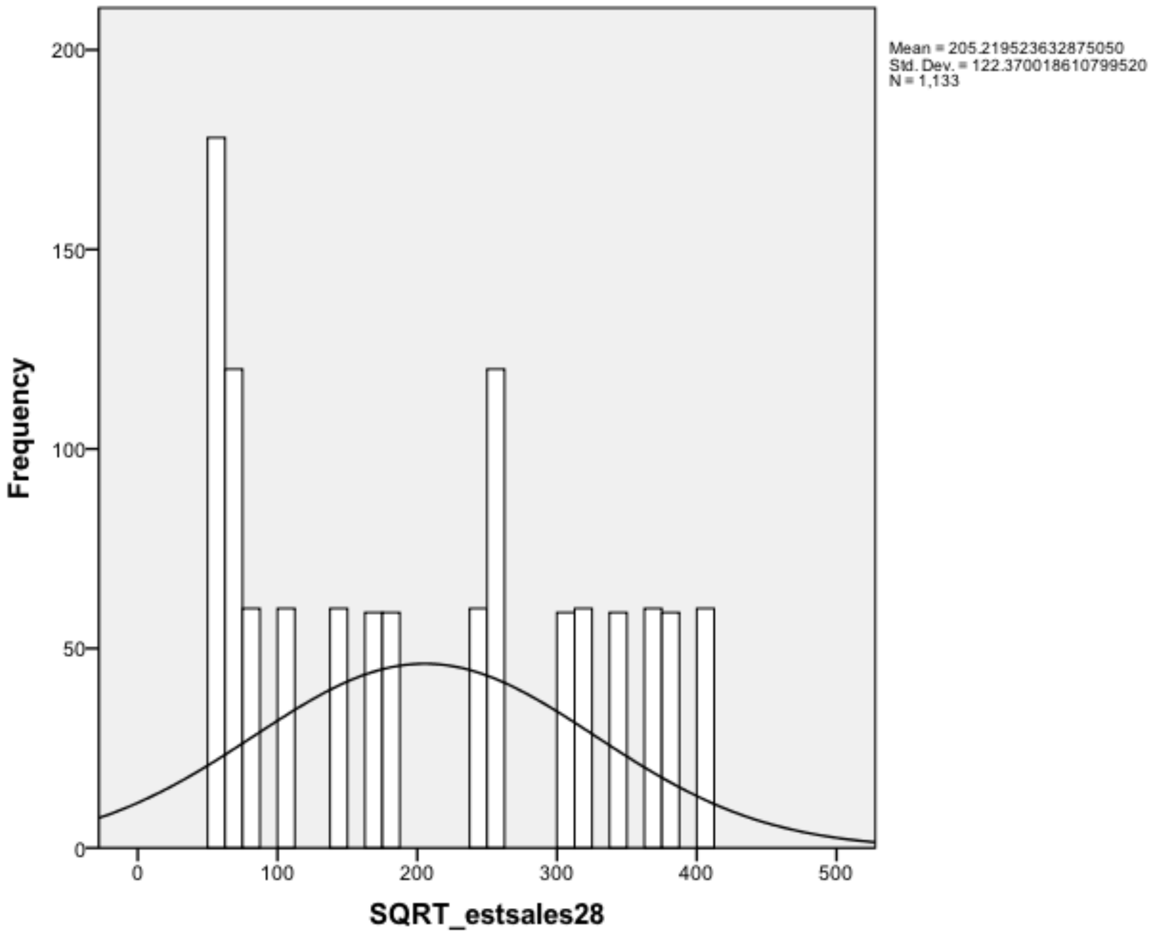
Appendix A1. Aggregated dataset - Distribution of errors for 'estsales28'



Appendix A2. Aggregated dataset - Distribution of errors for 'SQRT_estsales28'

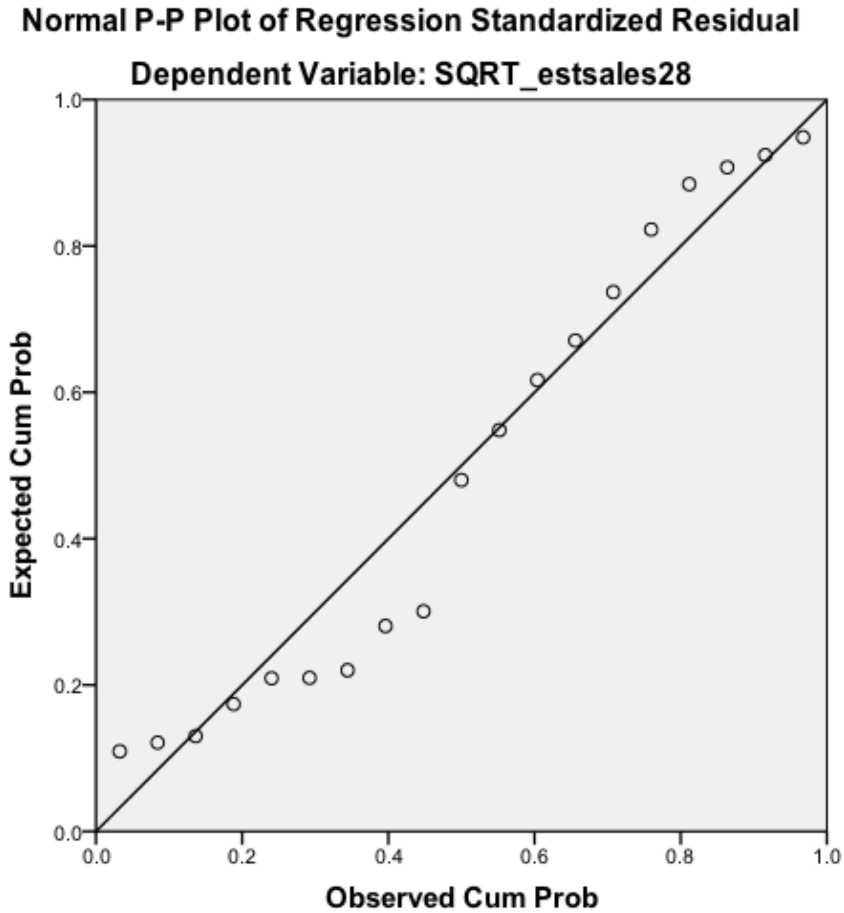


Appendix A3. Disaggregated dataset - Distribution of errors for 'estsales28'

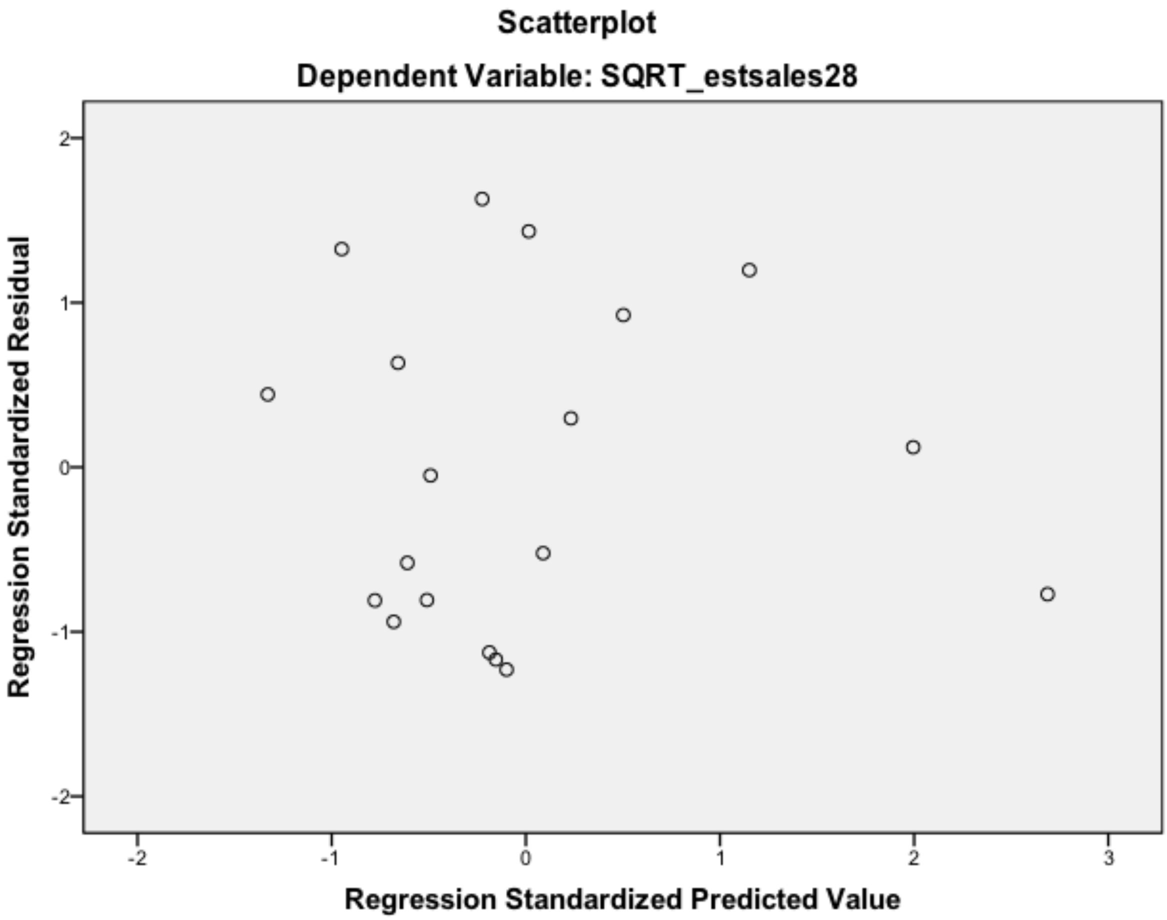


Appendix A4. Disaggregated dataset - Distribution of errors for 'SQRT_estsales28'

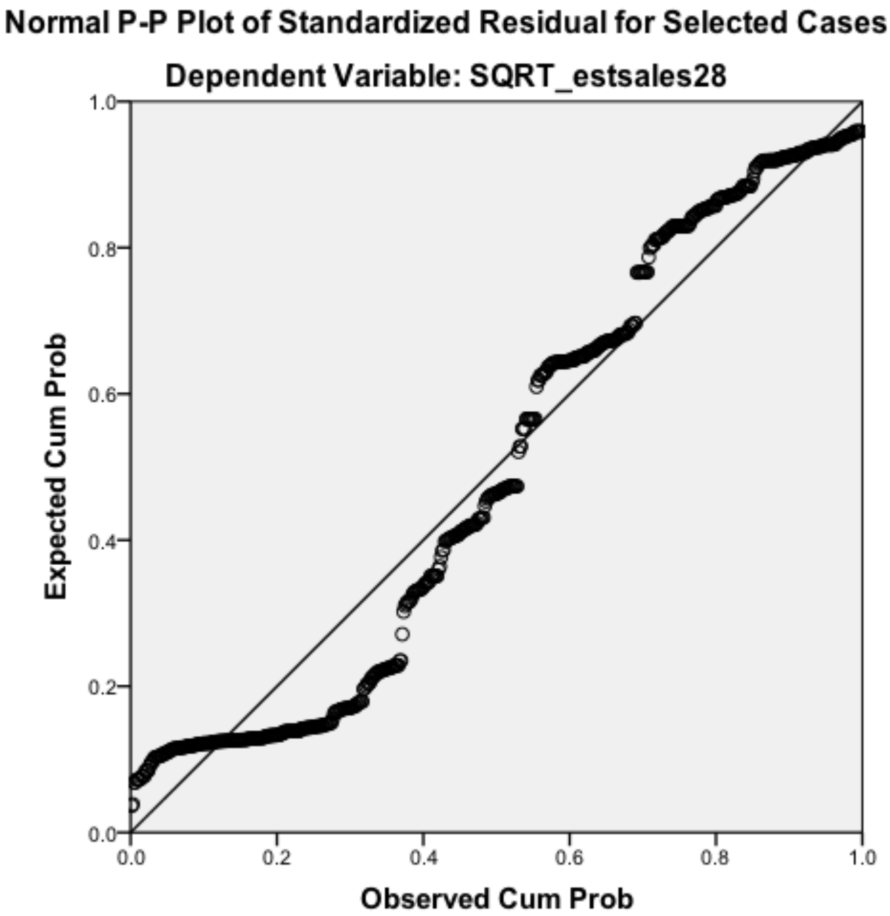
APPENDIX B



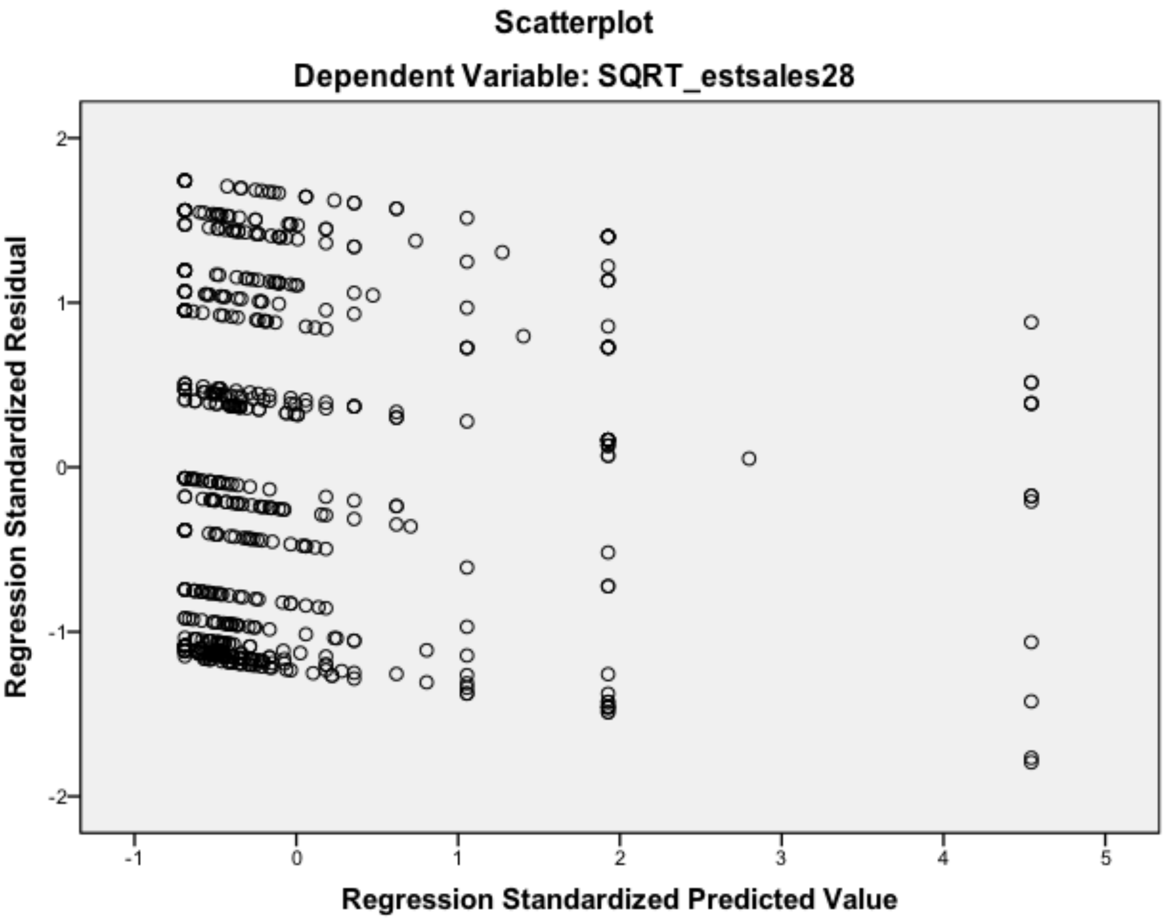
Appendix B1. H1a – Normality of errors assumption check for aggregated dataset



Appendix B2. H1a - Homoscedasticity assumption check for aggregated dataset

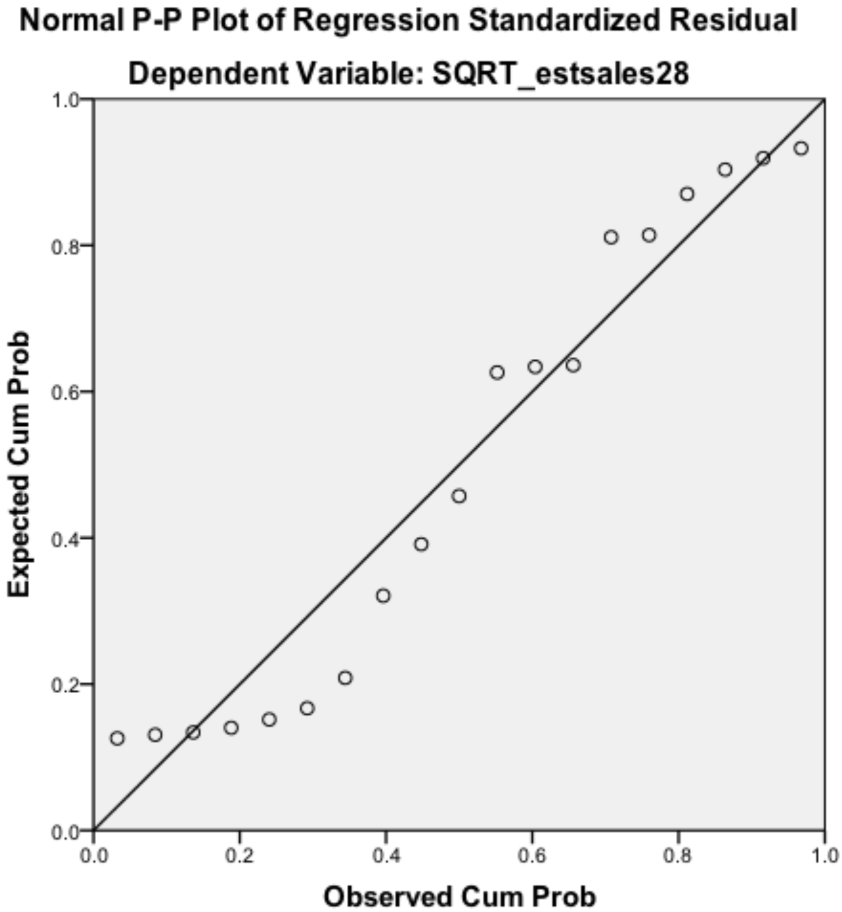


Appendix B3. H1a – Normality of errors assumption check for disaggregated dataset

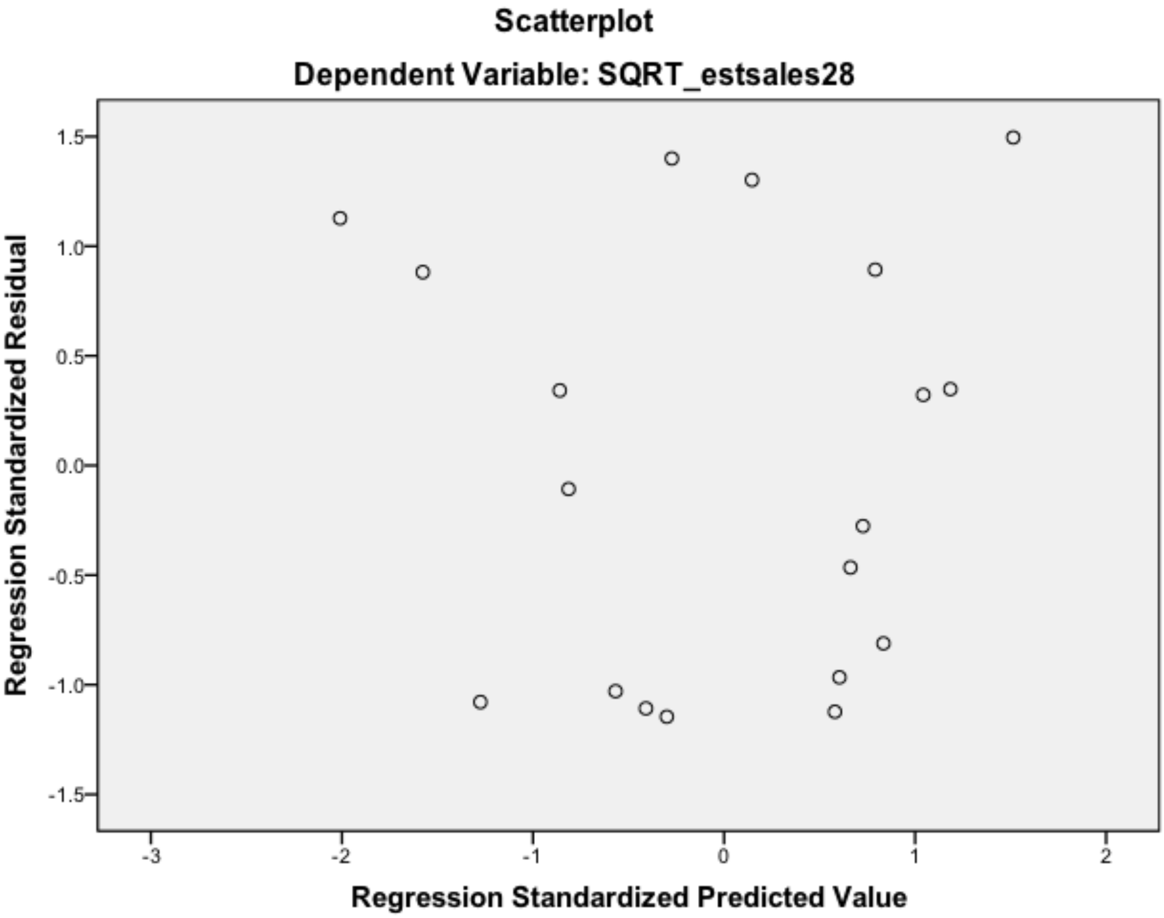


Appendix B4. H1a - Homoscedasticity assumption check for disaggregated dataset

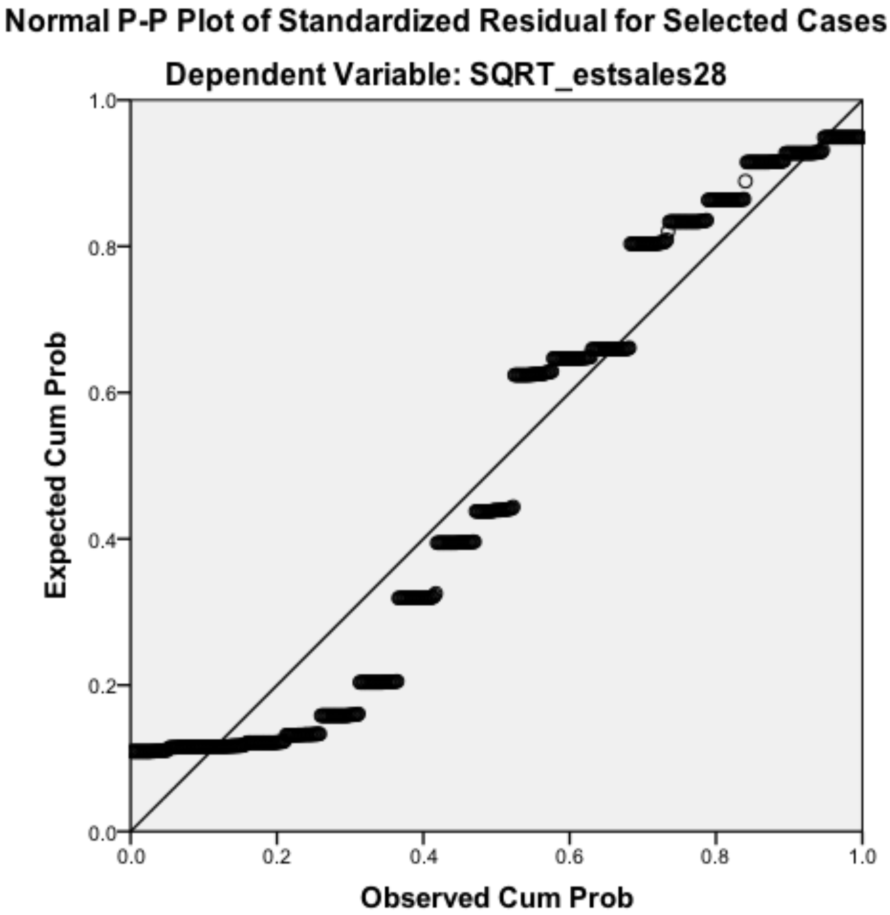
APPENDIX C



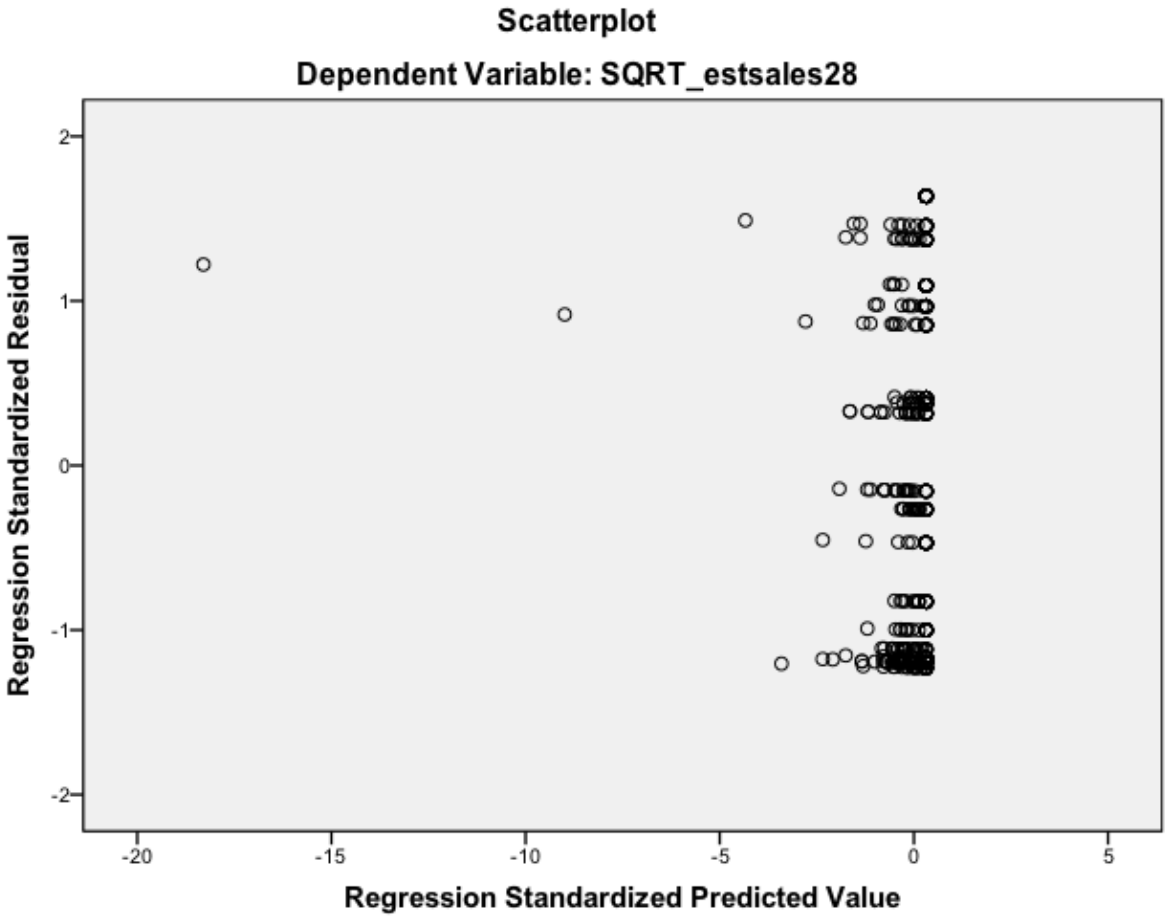
Appendix C1. H1b - Normality of errors check for aggregated dataset



Appendix C2. H1b - Homoscedasticity assumption check for aggregated dataset

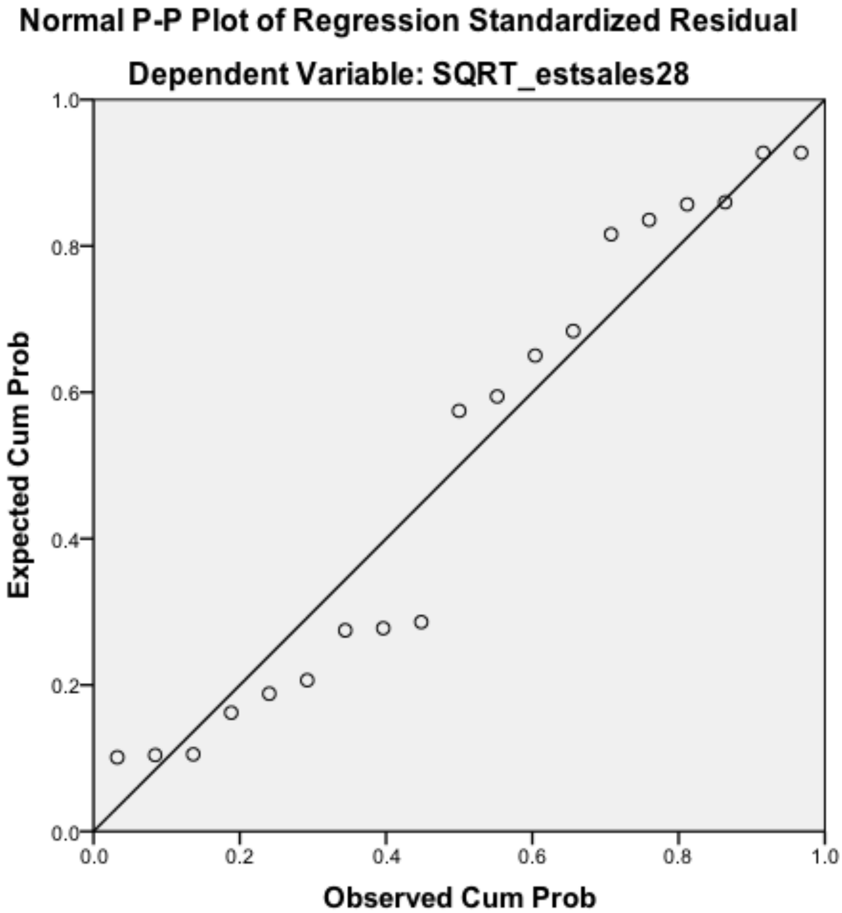


Appendix C3. H1b - Normality of errors check for disaggregated dataset

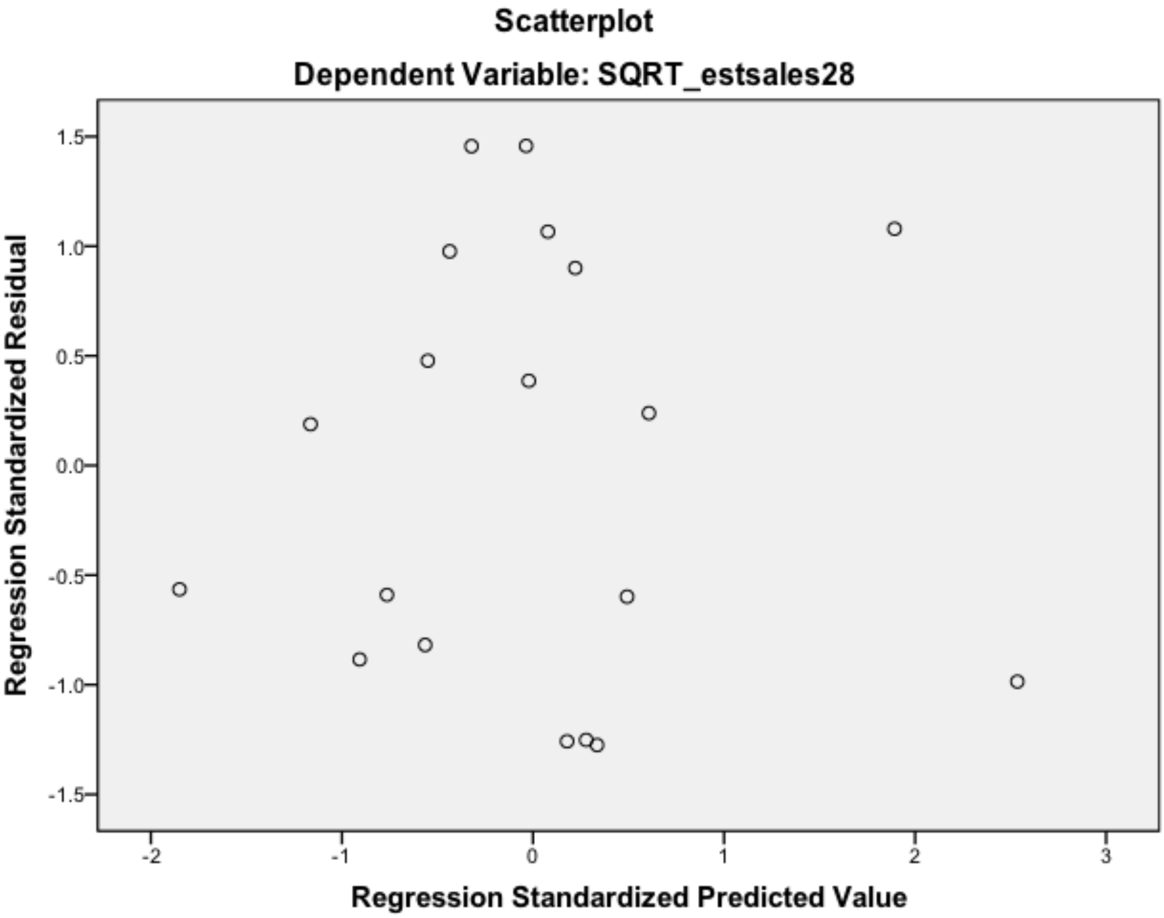


Appendix C4. H1b - Homoscedasticity assumption check for disaggregated dataset

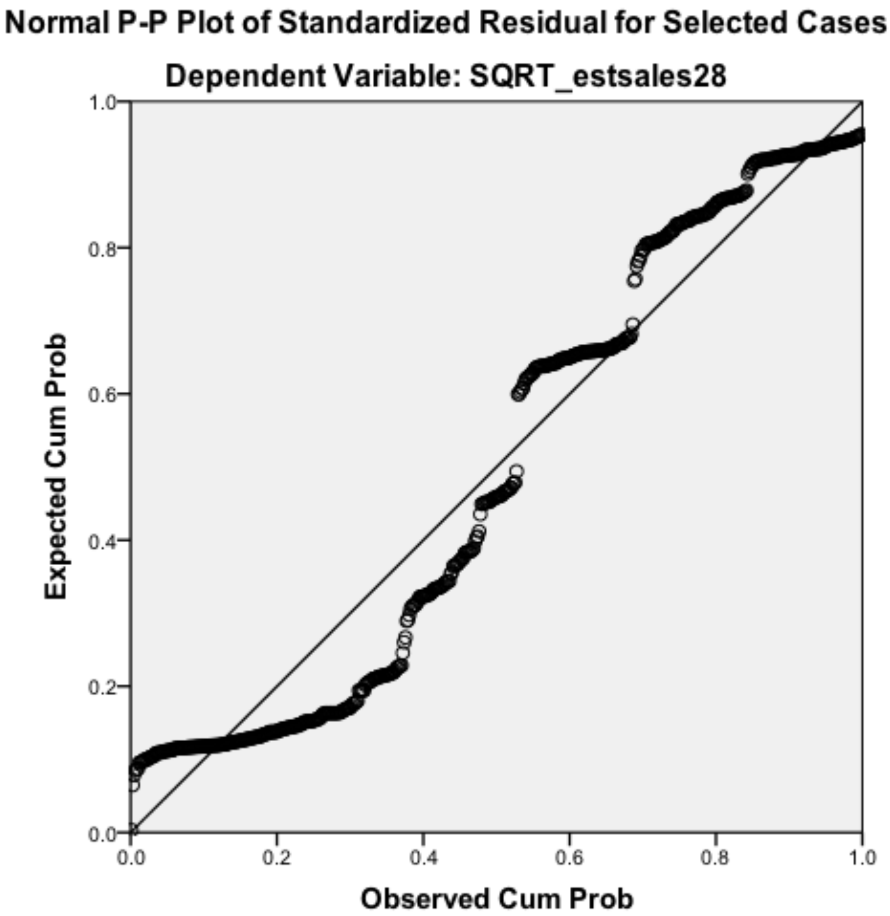
APPENDIX D



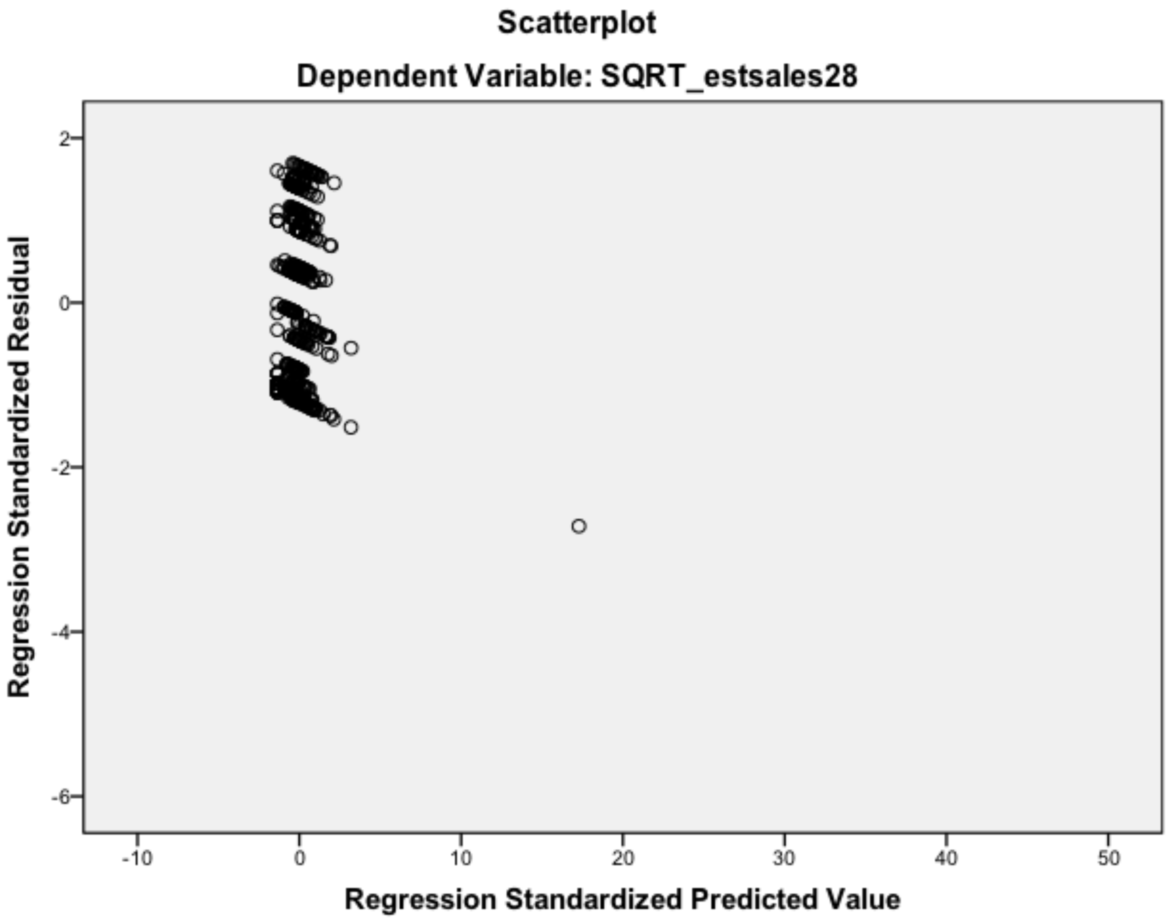
Appendix D1. H2a - Normality of errors assumption check for aggregated dataset



Appendix D2. H2a - Homoscedasticity assumption check for aggregated dataset

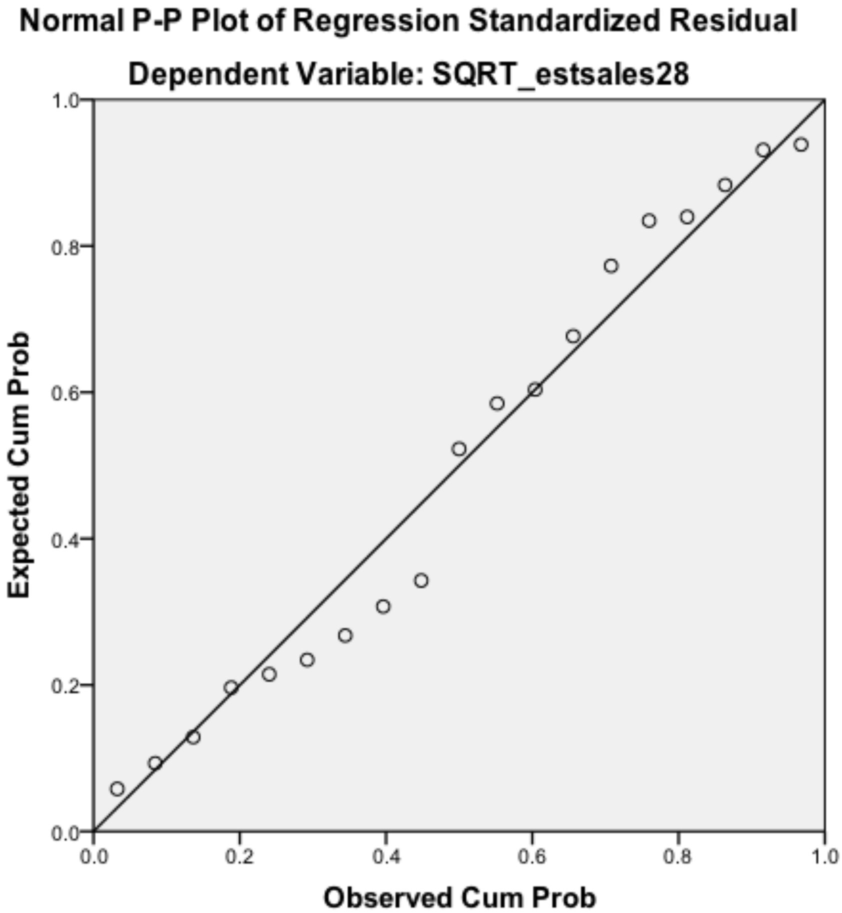


Appendix D3. H2a - Normality of errors assumption check for disaggregated dataset

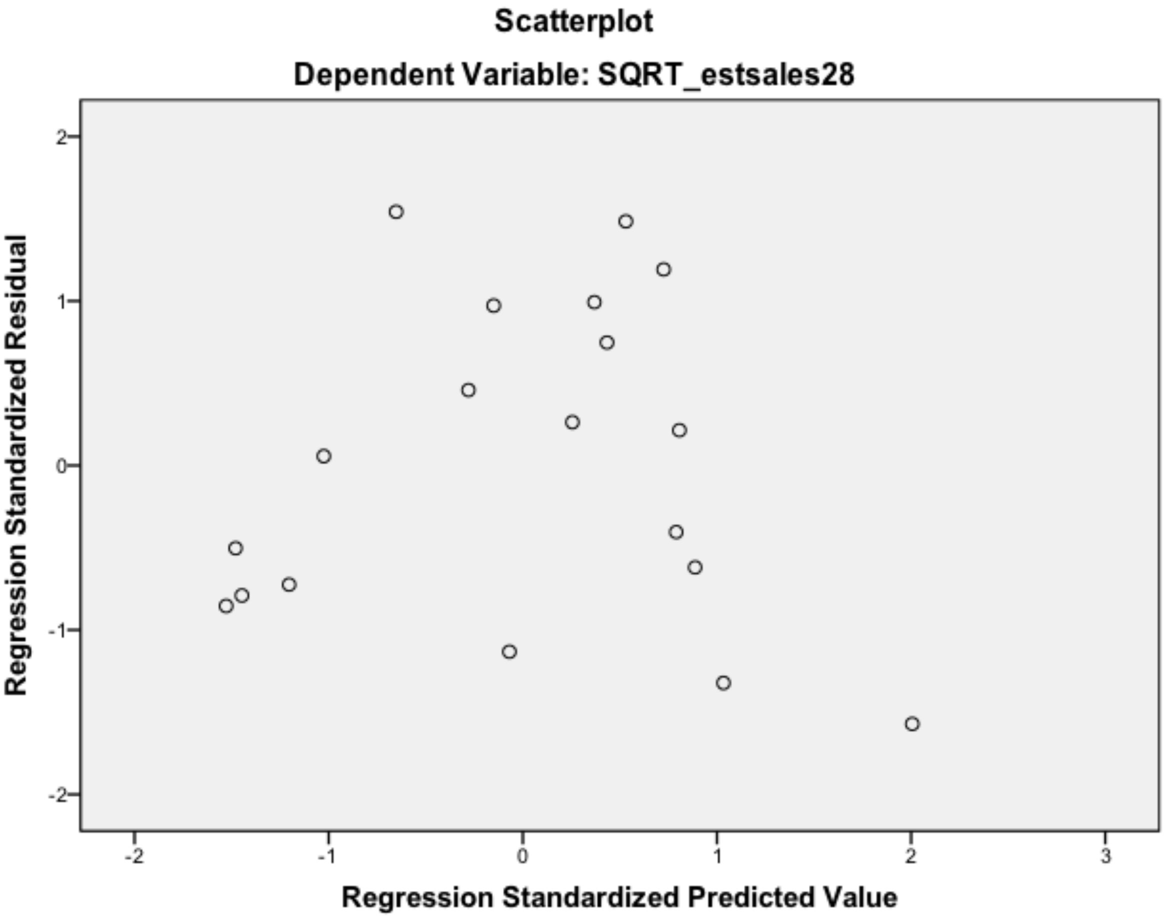


Appendix D4. Homoscedasticity assumption check for disaggregated dataset

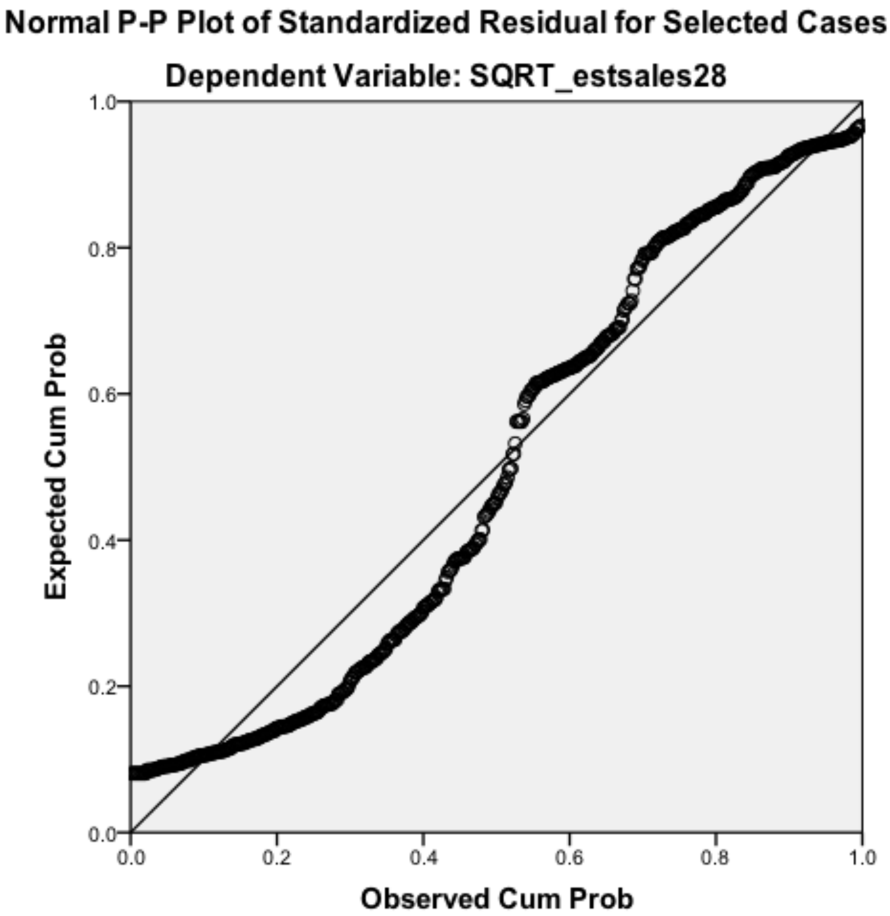
APPENDIX E



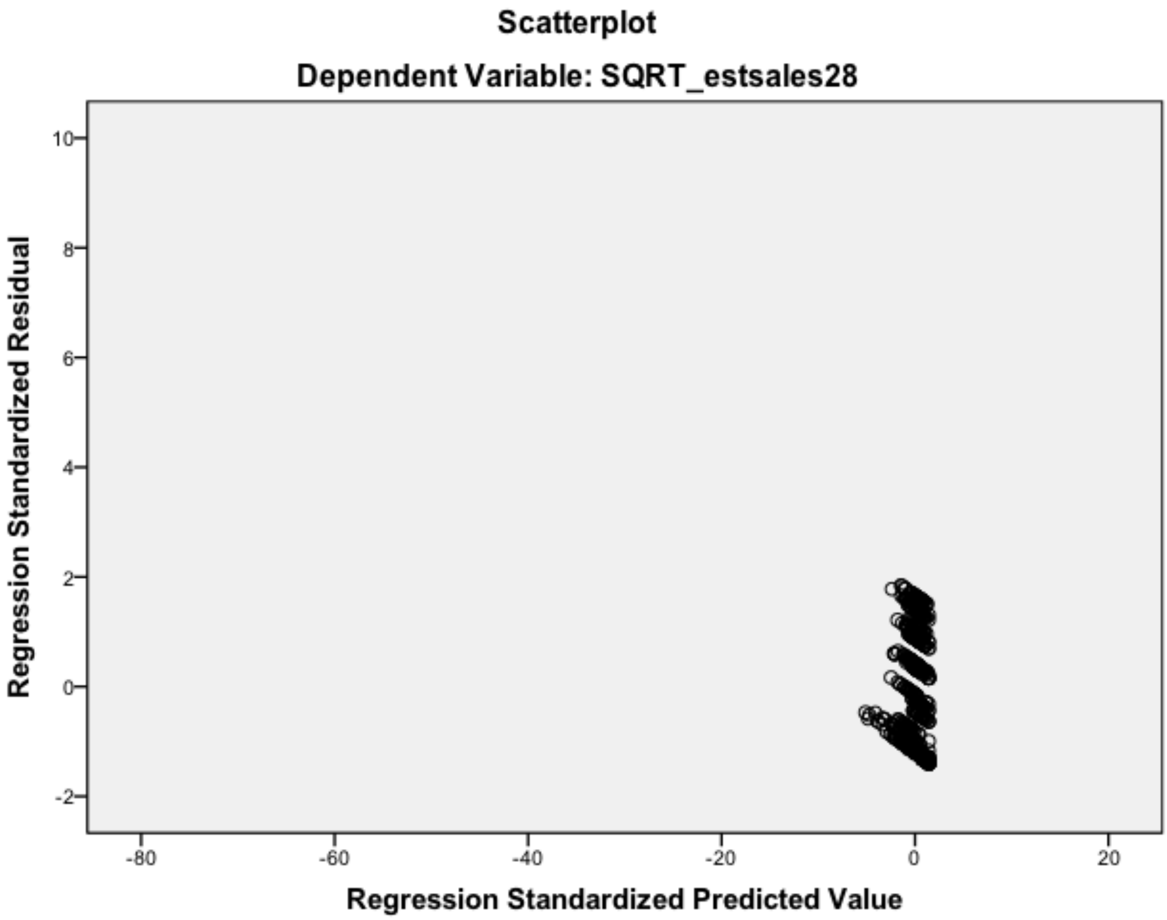
Appendix E1. H2b/H3a - Normality of errors assumption check for aggregated dataset



Appendix E2. H2b/H3a - Homoscedasticity assumption check for aggregated dataset

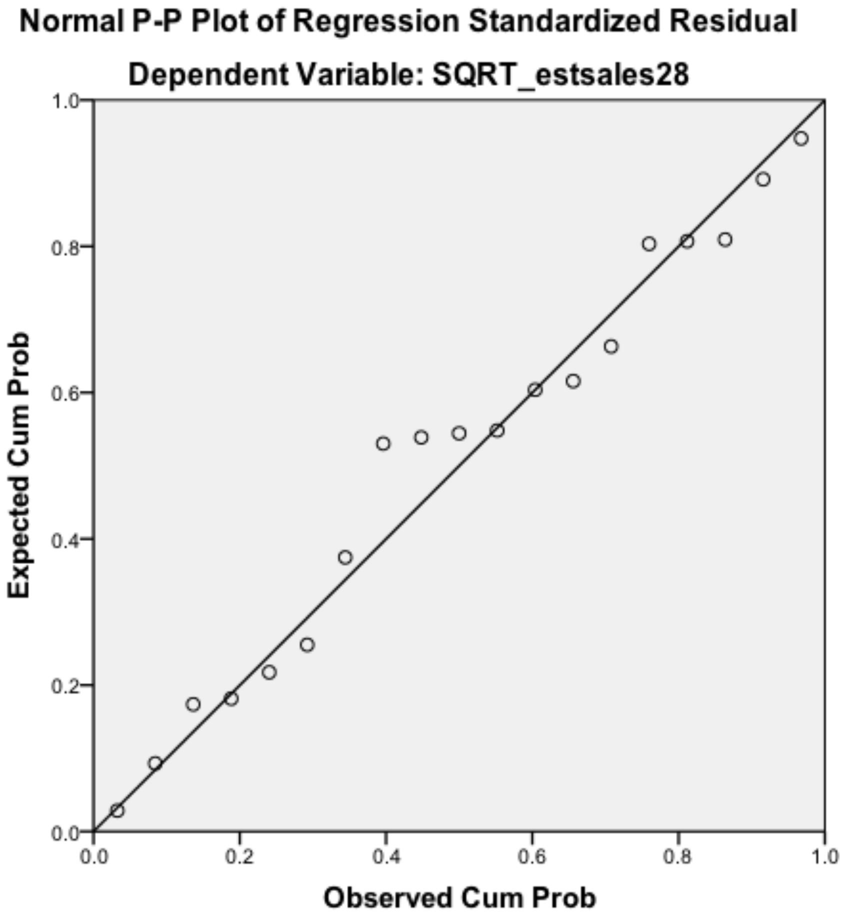


Appendix E3. H2b/H3a - Normality of errors assumption check for disaggregated dataset

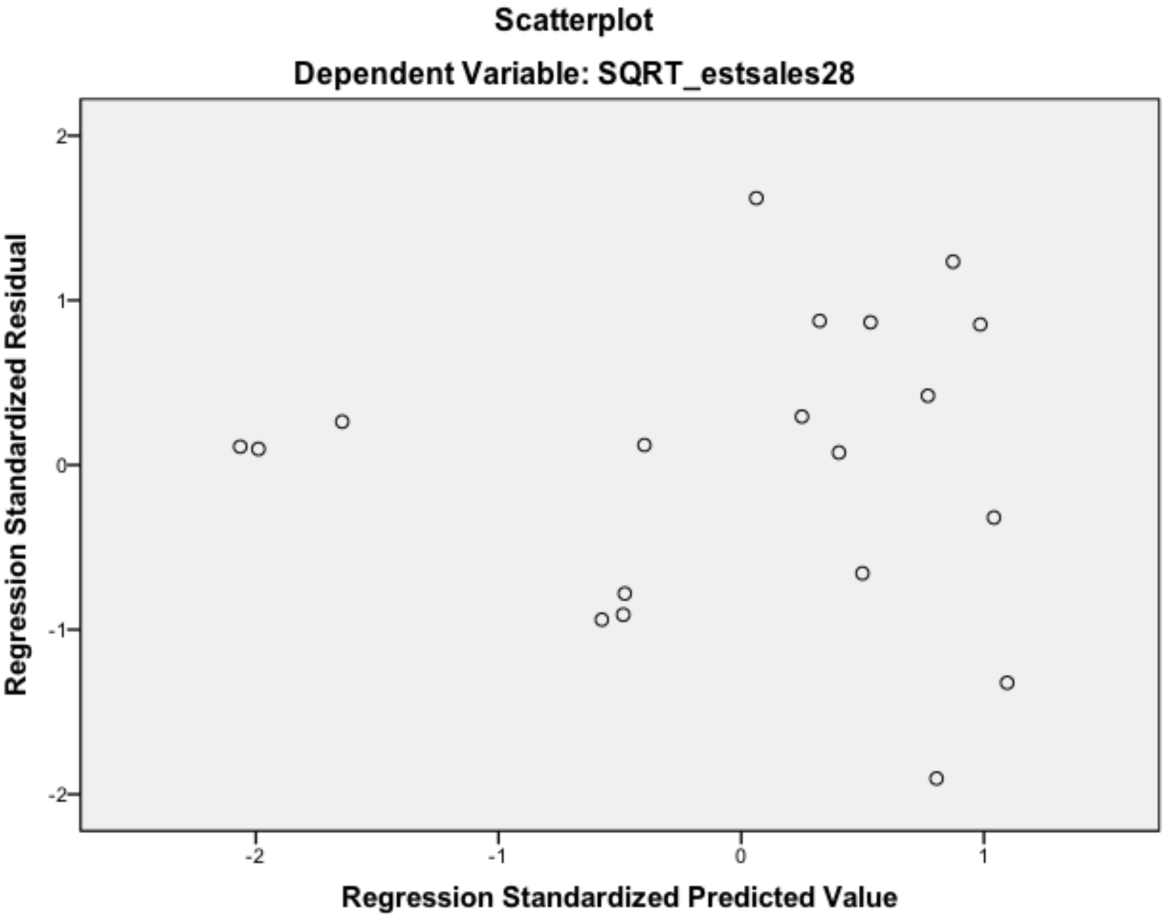


Appendix E4. H2b/H3a - Homoscedasticity assumption check for disaggregated dataset

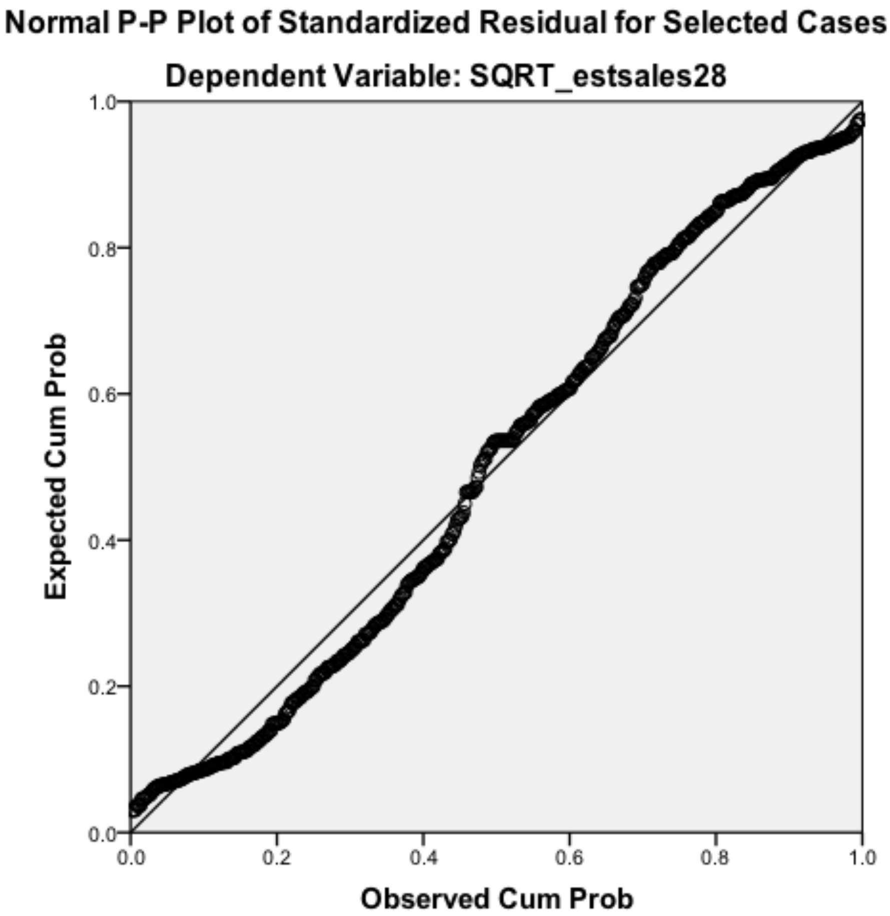
APPENDIX F



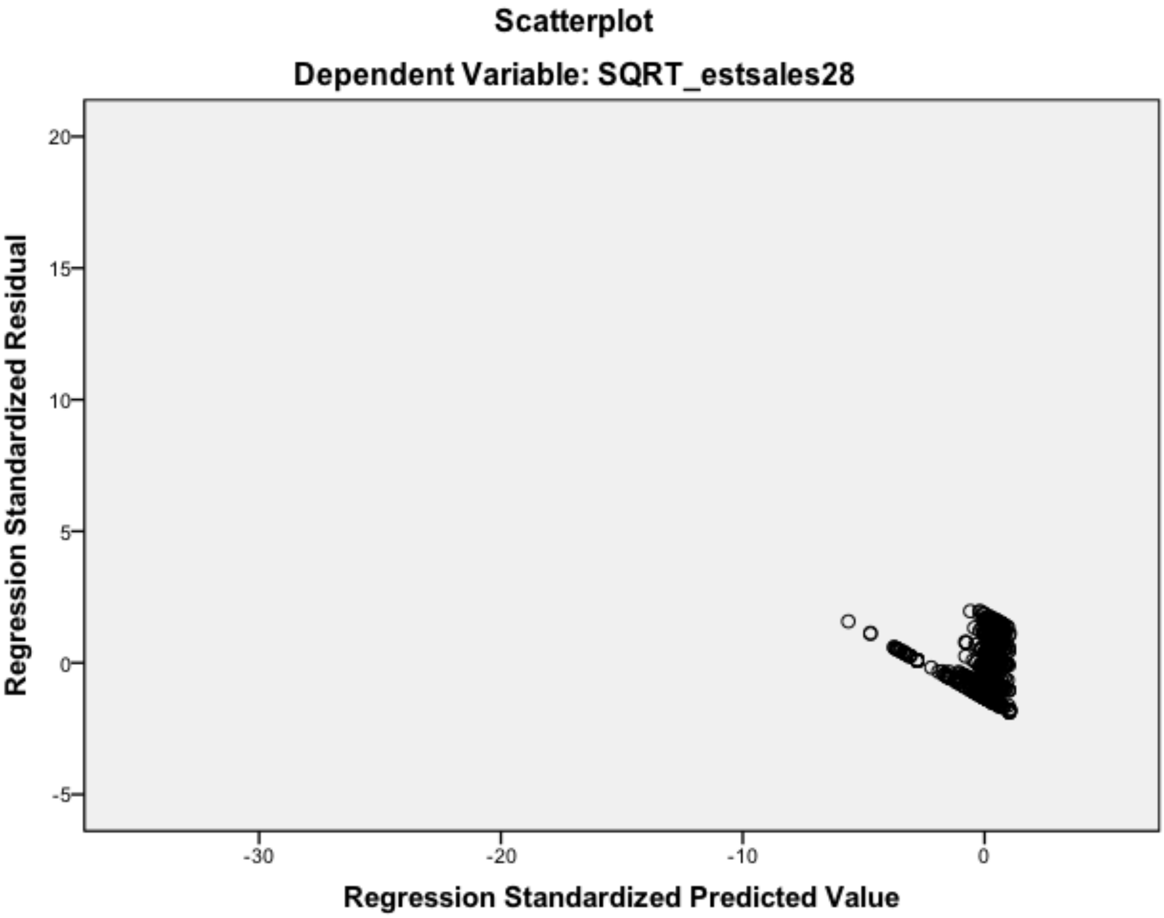
Appendix F1. H3b - Normality of errors assumption check for aggregated dataset



Appendix F2. H3b - Homoscedasticity assumption check for aggregated dataset

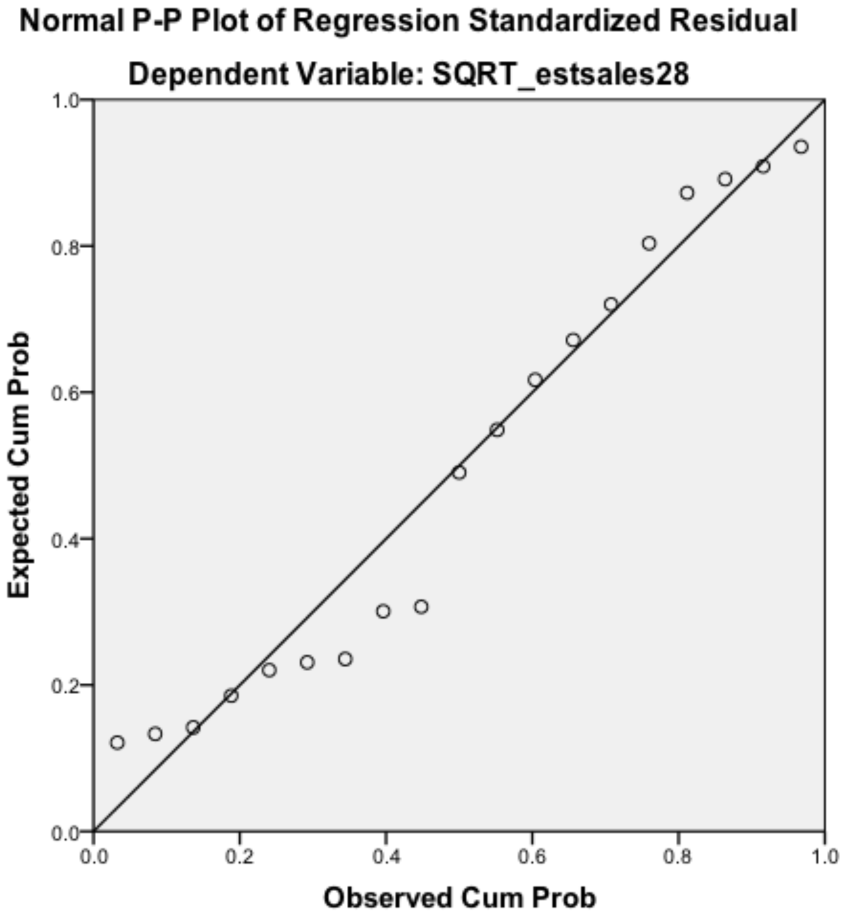


Appendix F3. H3b - Normality of errors assumption check for disaggregated dataset

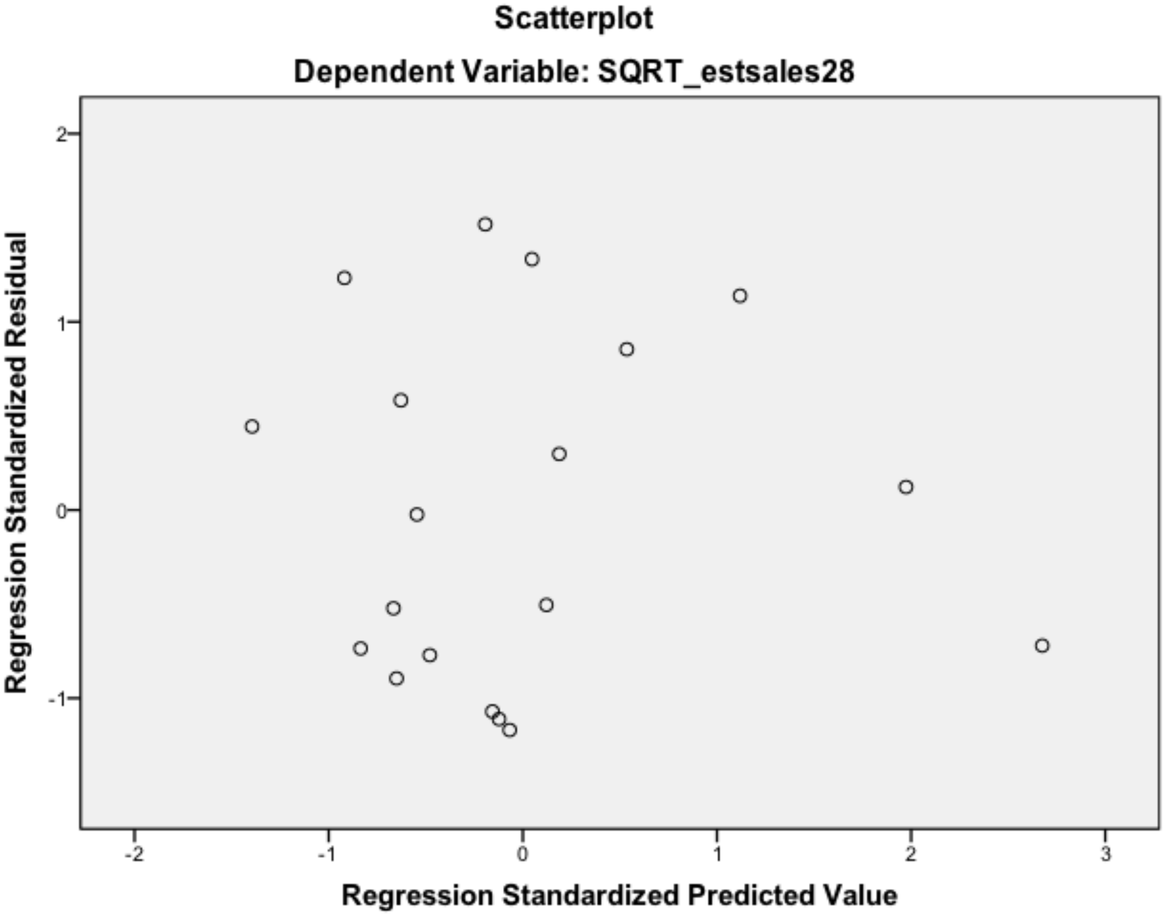


Appendix F4. H3b - Homoscedasticity assumption check for disaggregated dataset

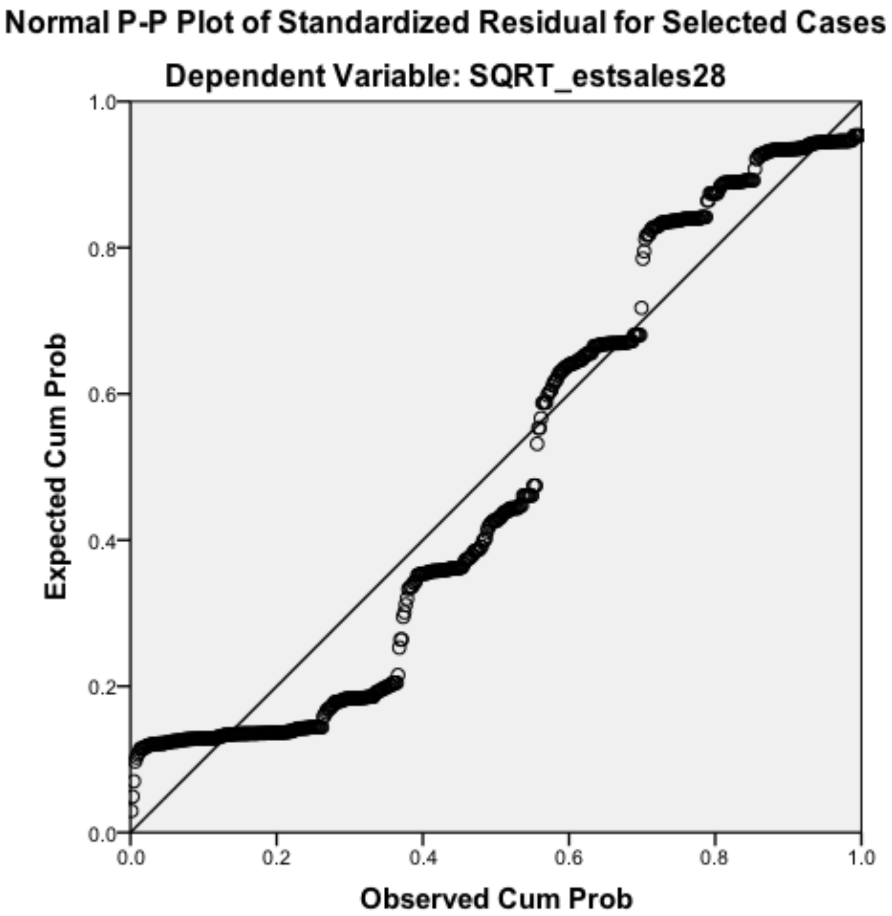
APPENDIX G



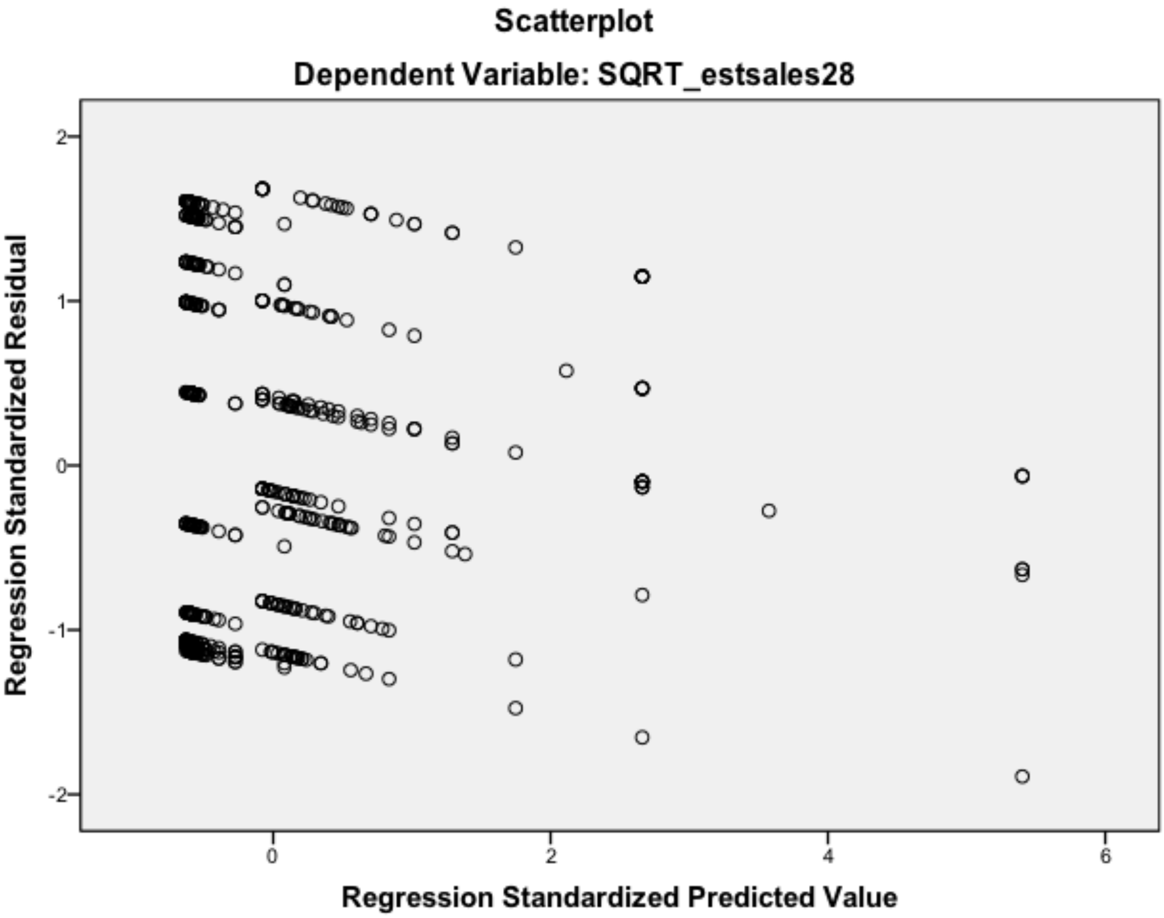
Appendix G1. H5 - Normality of errors assumption check for aggregated dataset



Appendix G2. H5 - Homoscedasticity assumption check for aggregated dataset

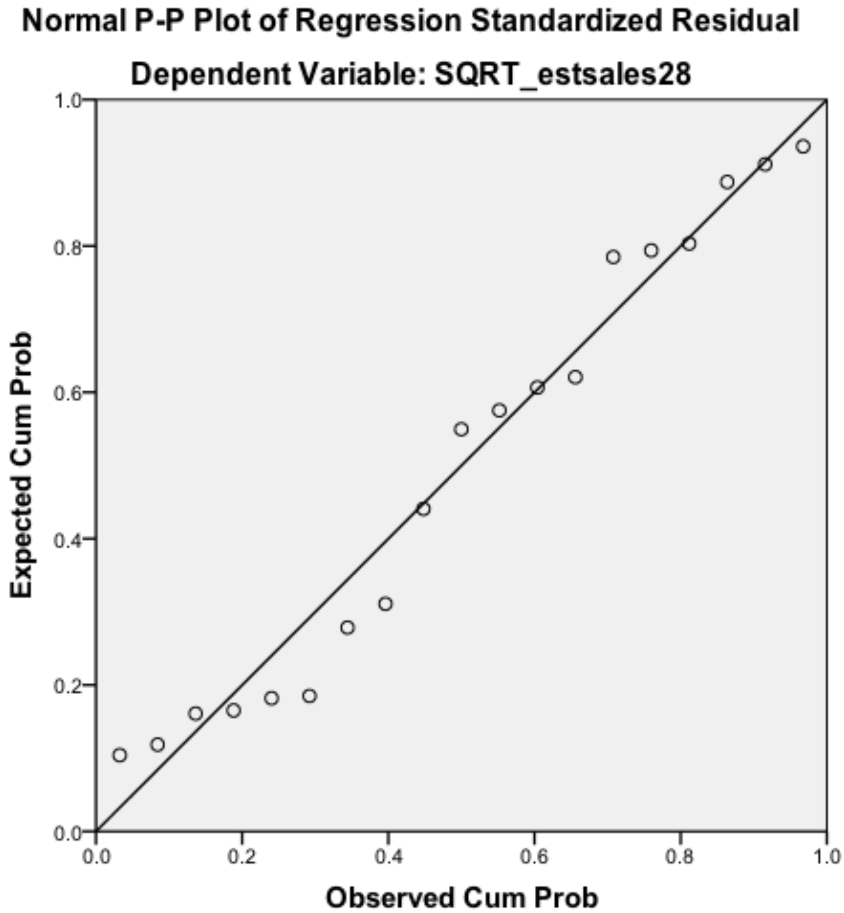


Appendix G3. H5 - Normality of errors assumption check for disaggregated dataset

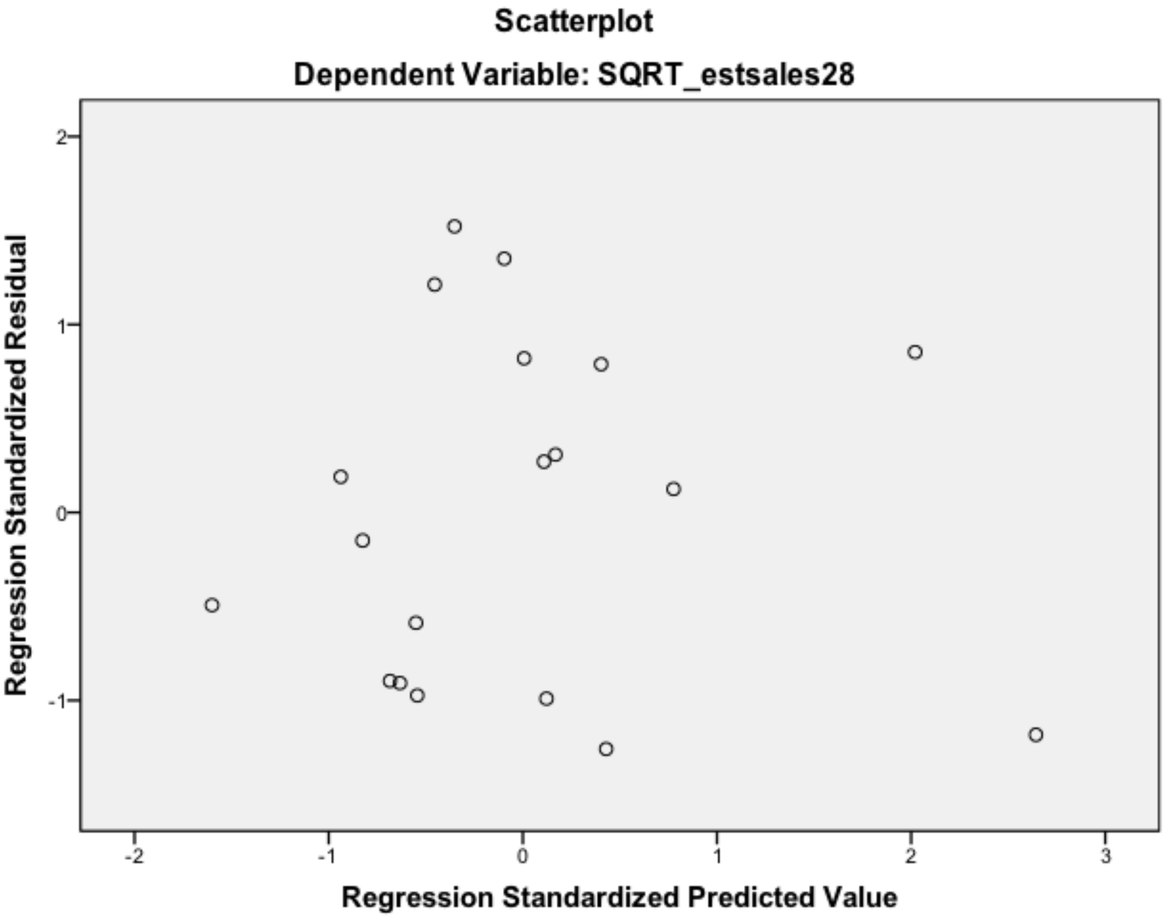


Appendix G4. H5 - Homoscedasticity assumption check for disaggregated dataset

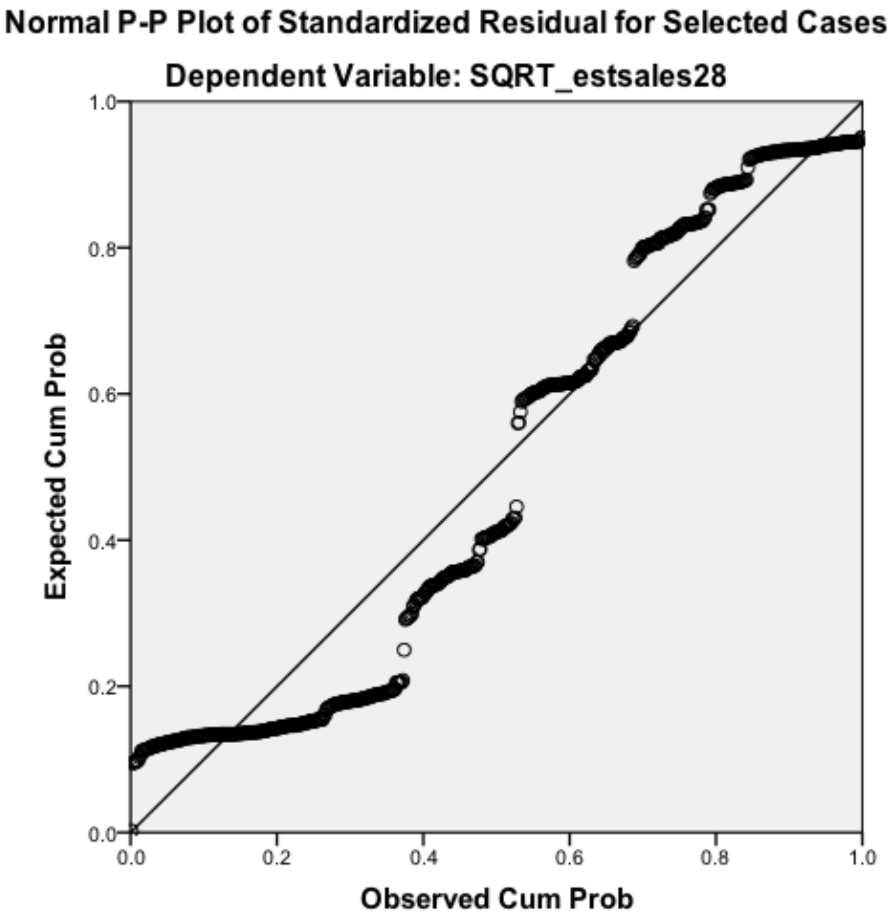
APPENDIX H



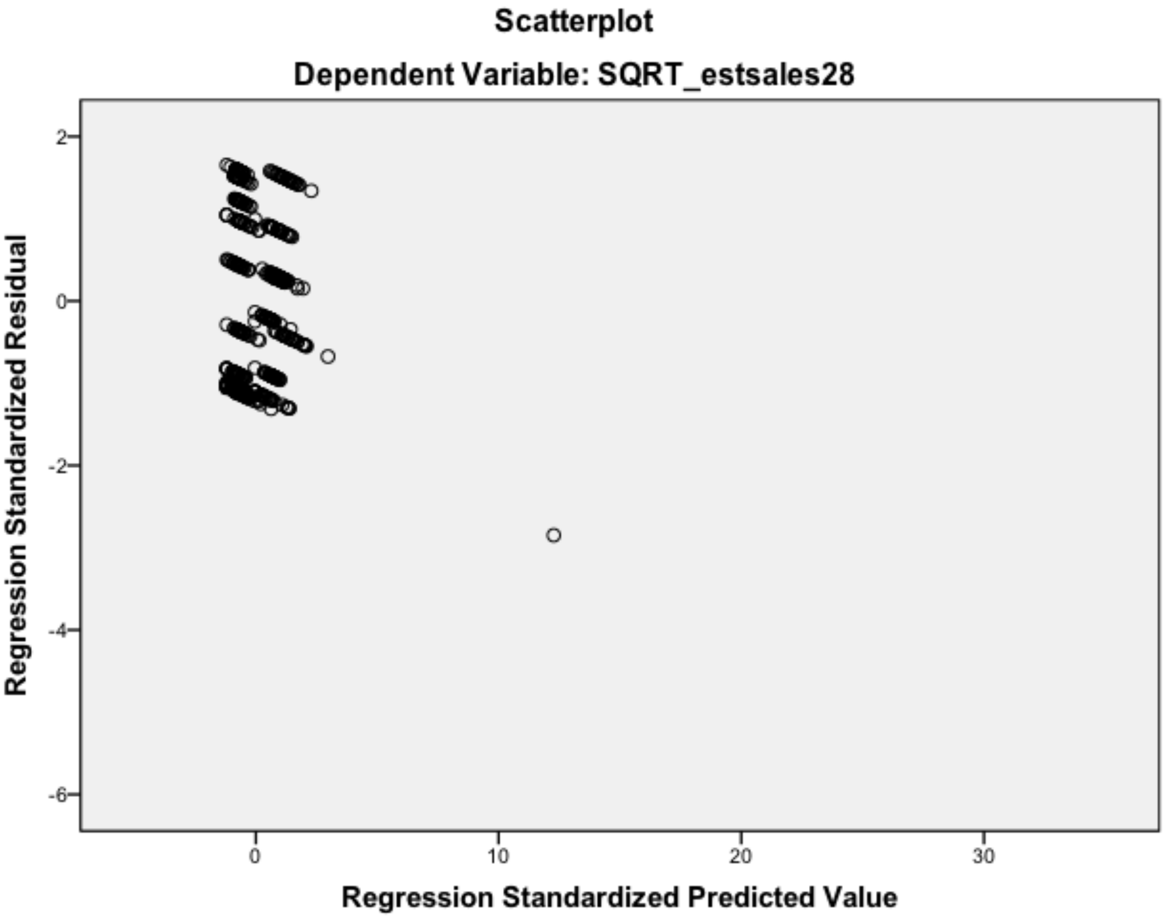
Appendix H1. H6 - Normality of errors assumption check for aggregated dataset



Appendix H2. H6 - Homoscedasticity assumption check for aggregated dataset



Appendix H3. H6 - Normality of errors assumption check for disaggregated dataset



Appendix H4. H6 - Homoscedasticity assumption check for disaggregated dataset

APPENDIX I

Memos on observations during the retrieval of expert reviews

1: it seems like books with lots of expert reviews tend to deal with / related to contemporary issues (e.g. American politics, feminism)
2: self-help books don't seem to receive many expert reviews
3: some reviews with different titles/sites have exactly the same content (probably newspapers all buy news 1 source e.g. Reuters?!)
4: have to check content manually for some titles to exclude brief news mentions of the books but do not really review its content. E.g. Kasich's "Two Paths: American Divided or United" - he's a presidential candidate so his name and book release pop up in all sort of political news, not just expert reviews
5: children books, novels, classics, religions books & self-help books don't get lots of expert reviews
6: memoirs (war, cancer) tend to get expert reviews
7: novels with popular movie adaptations can become bestsellers and so get covered by expert reviews
8: some award-winning novels (e.g. the girl who drank the moon) —> expert reviews
9: topics about racism (e.g. black life matters) receive media attention
10: many books of similar genre/topic are reviewed together
11: fictions (e.g. heroes, magic) don't get that many reviews