Brand Equity and Country-of-Origin Effects on Consumer Evaluations: Leveraging Data to Explore the Role of Country Image and Brand Familiarity on *Amazon* Product Reviews

Kathy Tian
The University of Illinois at Urbana-Champaign

Mike Yao
The University of Illinois at Urbana-Champaign

As globalization accelerates transatlantic trade, both developing and developed countries are experiencing an entry of more foreign products simultaneous to rising exports. Against the backdrop of a progressively fluid marketplace, the question of how a product's country-of-origin (COO) influences consumer judgement has propelled in significance. Although the consequences of country-of-origin effects have received robust attention in extant literature, few studies have reproduced the results on e-commerce platforms, which are rapidly replacing conventional brick-and-mortar retailers. In this study, the authors explore the extent to which origin effects influence consumer perceptions on Amazon, the world's largest online retailer, by investigating (1) how Amazon consumers rate products based on country-of-origin cues such as level of economic prosperity, (2) if brands from developed countries are perceived as more familiar than those from developing countries, and (3) if brand familiarity moderates country-of-origin effects. The findings from this study offer significant implications to both the domains of marketing and consumer studies as well as to marketing practitioners.

Keywords: Country-of-origin, Brand, Consumption, e-Commerce

INTRODUCTION

Companies rely on brands to differentiate their products in an increasingly saturated consumer goods market (Bristow et al., 2002). Likewise, in service industries, brands represent a heuristic for consumers; i.e., consumers leverage brand information to make snap decisions on which service or product is optimal for their needs. As labels accelerate information processing, hold social and emotional values to their consumers, and enhance perceived utility, branding invariably influences a company's overall profit margins (Kotler & Gertner, 2002). As such, branding reflects an integral component of a product's success, particularly in the international market where competition is intensifying due to market fluidity.

Considering this increasing fluidity in both international and domestic markets, country image, as a component of brand image, has propelled in significance. Accordingly, the country from which a given brand originates influences consumer evaluations, perceptions, and subsequent purchase intentions

(Koubaa et al., 2013, Verlegh & Steenkamp, 1999). The country-of-origin effect (COO) describes this phenomenon. COO refers to the situation that occurs when consumers' baseline evaluation of a product, service, or brand are altered owing to an association between that particular entity and a given location (Andéhn et al., 2016). A brand's country-of- origin could greatly influence consumers' perceptions of that brand or its related products. This is particularly true when limited cues concerning the quality of an item are available for consumers to evaluate (Olga et al., 2017). COO has received notable attention in the international marketing literature and there exists robust literature on both the influential role it plays in consumer perceptions as well as situations where its relevance is attenuated (Usunier, 2011).

However, although much of the present literature on the country-of-origin effect indicate consumers tend to prefer products and brands from wealthy nations (Verlegh & Steenkamp, 1999) and that consumers also find brands from certain countries more familiar than others (Koschate-Fischer et al., 2012), there is still limited research concerning how well these findings can be reproduced in an online setting. Clearly, the COO effect holds significant implications to both scholars and practitioners alike, yet there is an evident lack of research concerning the relationship between a product's country-of-origin and customer evaluations in an e-commerce setting. In this study, we aim to address this gap by reproducing past studies; as such, we follow a natural experiment design and use customer data from Amazon, a leading global e-commerce retailer. In particular, we examine how the effects of economic prosperity of a given country and familiarity with a brand influences consumer evaluation on Amazon.

CONCEPTUAL FRAMEWORK

Country-of-origin Effects and Moderators

Over the past five decades, a robust body of literature on country-of-origin effects has accumulated, with scholars suggesting that origin effects are a central component of international marketing (see Tan & Farley, 1987; Kotler & Gertner, 2002; Magnusson et al., 2011). Interestingly, however, research further suggests that origin effects on product evaluations are not uniform; that is, the degree to which a product's country-of-origin impacts consumer judgements appears largely context-dependent (O'Shaughnessy & O'Shaughnessy, 2000). A brand's country-of-origin, for instance, may bolster or undermine the brand's perceived value depending on product category. For example, consumers may generally perceive items originating from Colombia as less desirable, but when considering Colombian coffee, the same consumer's image of Colombian products may elicit ideas of rich fields yielding rich coffee (Kotler & Gertner, 2002; Andéhn et al., 2016).

The COO effect, which describes how consumer evaluations of a product may differ based on the country the product was exported from, has also been found to correlate with a country's perceived level of economic development (Verlegh & Steenkamp, 1999; Martin et al., 2011; Andéhn et al., 2016). As an example, products from highly developed countries with a long-standing history of quality production, such as Japan and Germany, are perceived more favorably than similar products from developing nations (Martin et al., 2011). Further, in their meta-analysis on country-of-origin effects, Verlegh and Steenkamp (1999) conclude that differences in economic development underlie origin effects: products from economically wealthy nations tend to outperform products from countries in a transitionary state. Taken as a whole, these studies indicate that the magnitude of country-of-origin effects are generally contingent on factors or situations related to the product or brand (e.g., product category, economic development, etc.) (Verlegh & Steenkamp, 1999; Andéhn et al., 2016).

Brand Familiarity and Country-of-origin

Research on consumer brand and product perceptions often yield conflicting results, with some studies suggesting consumers find emotional advertising appeals more persuasively, whereas others argue the opposite (Lau-Gesk et al., 2009). Evidently, consumer brand perception is a labyrinthine topic; for example, consumers may not always know what they want and attitudes toward products are not always easily measured. As such, the effects of country-of-origin alone cannot fully capture the intricacies underlying what informs consumer perceptions. Studies indicate, however, that brand familiarity may explain the variances in consumer attitudes toward brands or products (Roy & Bagdare, 2015; Rhee & Jung, 2018). For instance, in their study on advertising effectiveness and brand familiarity, Rhee and Jung (2018) find a consumer's familiarity with a brand significantly moderates the strength of the relationship between ad attitude and brand attitude. Similarly, research suggests brand familiarity also moderates the effect of a celebrity endorser on consumer attitudes toward a product (Roy & Bagdare, 2015).

Brand equity, or familiarity, significantly influences the relationship between a consumer and their perceptions toward a given product, and this relationship extends to the context of origin effects as well. As an example, studies on brand equity and COO show that consumers' familiarity toward a brand and its country alliance moderates overall evaluations and perceptions (Kristensen et al., 2014). Further, previous investigation concerning the effects of brand familiarity on COO suggests knowledge of a brand does not always produce positive perceptions. Specifically, familiarity with a brand actually results in a negative moderating influence on the COO effect in high-involvement settings. Additionally, past studies indicate that country-of-origin effects are amplified when a brand is relatively unknown as consumers may make assumptions about that brand based on its country-of-origin (Koschate-Fischer et al., 2012). More evidence, however, is needed in the research community to establish and reproduce the moderating effect of brand equity on COO in online shopping environments.

Country-of-origin Effects in E-commerce

There remains relatively limited research on how the country-of-origin effects replicate to e-commerce platforms, which are becoming increasingly common as an alternative to traditional brick-and-mortar retail stores (Yan et al., 2018). As globalization and natural country-specific competitive advantages continue to accelerate transnational trade, the importance of understanding COO and its influence on consumer perceptions of foreign goods rises simultaneously. However, with new technologies and the growing ease of online shopping via major e- commerce platforms, it is important to question how valid COO effects are in the context of digital consumerism. As such, we aim to address this concern by replicating past findings on the origin effects in the context of e-commerce platforms. In this study, we examine consumer data from Amazon.com, the world's largest online shopping retailer, to evaluate the extent to which a product or brand's country-of-origin is predictive of subsequent consumer ratings.

Specifically, based on previous research concerning the role of country-image (e.g., economic wealth) and brand equity on the country-of-origin effect, we are interested in understanding 1) how a brand's country-of-origin impacts the valence of consumers' reviews on Amazon, 2) how a brand's country of origin is related to brand familiarity, and 3) whether brand familiarity influences the magnitude of the country-of-origin effect on online shopping sites? Considering these three questions, we propose the following hypotheses: H1) Amazon consumers will rate products from developed (developing) countries more (less) favorably than products from developing (developed) countries; H2) Amazon consumers will rate highly familiar (unfamiliar) brands more (less) favorably than unfamiliar (familiar) brands; and H3) Brands from developing nations will see a steady increase in ratings over time.

Hypotheses

***H1**: Amazon consumers will rate products from developed (developing) countries more (less) favorably than products from developing (developed) countries.*

***H2**: Amazon consumers will rate highly familiar (unfamiliar) brands more (less) favorably than unfamiliar (familiar) brands.*

***H3**: Brands from developing nations will see a steady increase in ratings over time.*

METHODOLOGY

Study 1 – The Country-of-origin Effects on Amazon Ratings

Design

Study 1 follows a natural experiment pattern as this study is an analysis of consumer behavior on Amazon. Data on what products consumers purchased as well as the follow-up ratings that consumers attributed to those products were archived online. To conduct this study, the authors retrieved Amazon data that was generously archived by Julian McAuley from the University of California, San Diego. In study 1, we explore how a product's country-of-origin, particularly the economic development of that country, predicts how favorable consumers found the product to be. Because Amazon customers often leave product reviews, using a scale from 1 star (low) to 5 stars (high), Amazon provides a unique natural experiment setting to examine the effects of country-of-origin on customer perceptions of the product or brand.

Given these conditions, study 1 follows a 2 by 2 (country-of-origin: developed vs. developing x consumer rating: low vs. high) natural design. We theorize a country's economic status will predict consumer evaluations of their product purchases; explicitly, we hypothesize that a more economically developed country will be associated with higher average reviews on Amazon compared with a developing country.

Participants

Participants were not actively recruited for this research as it follows a natural experiment design. Instead, participants comprised Amazon customers who had purchased a product from Amazon and left a rating on a scale of 1 to 5. In total, we analyzed data from 1,003,650 Amazon consumers, prior to data cleansing (electronics and fashion data combined).

Measures

Economic Development

The first independent variable we explore is country-of-origin. Particularly, in this instance, we explore how economic wealth (developed or developing economy) may be predictive of consequent customer ratings on the Amazon e-commerce retail site. We classify a country as developed or developing based on reports from the United Nations, an international organization committed to supporting international peace.

Consumer Rating

A natural real-world method is employed to measure the dependent variable, customer perceptions of the product or brand. As consumers are asked to leave a rating out of 5 stars on Amazon, with 1 star reflecting a low rating and 5 stars reflecting a high rating, we are easily able to gauge whether consumers held positive, negative, or neutral, perceptions of their purchased item.

Materials

Again, as study 1 follows a natural experiment design, we use materials archived online. We retrieve Amazon data on the purchased items including the date and time of purchase, the price, and the brand (as the indicator of country-of-origin). Additionally, we only use datasets containing ratings for the products so we are naturally able to obtain ratings through the Amazon data archive.

Procedure

Amazon customers were not actively recruited in study 1; however, their aggregate consumption patterns on the e-commerce site were available for analysis. As such, study 1 involved over 1,003,650 data points reflecting Amazon customers' shopping patterns on the platform. The aggregate data include two product categories: electronic goods (e.g., printers, computers, etc.), which contained a total of 954,251 products prior to data cleansing, and fashion items (e.g., watches, jewelry, clothing, etc.), which contained a total of 49,399 data points prior to cleansing. In this study, only dense datasets, or data that have been reduced to contain five core information points such that all users and items have a minimum of five reviews, were used. Using dense sets containing a baseline minimum of five reviews per product ensures that all items will correspond to a customer rating and there will be more than one rating per item, yielding higher quality data that is less prone to outliers.

Electronics

For the electronics category, there were a total of 954,251 data points with 3,525 unique brand names. Prior to data exploration, however, data cleansing was performed to rule out potential confounding variables. For instance, to ensure roughly homogeneous characteristics for the electronics category, only brands occurring 500 times or more were used, reducing the number of brands to 287. Additionally, to guarantee that each brand had multiple products, only companies offering more than 20 products were included in the data, resulting in 207 unique electronics brands. Finally, the country of origin for each of these unique electronic brand names was manually coded. Unfortunately, some of the brands extracted from data cleansing had no discernable country-of-origin and, thus, were dropped from the set. After manually coding all countries of origins for the remaining brands, a total of 100 unique instances were left.

These brands, along with each corresponding country-of-origin, were then appended back into the original electronics dataset. All items not included in the appended brand list were subsequently dropped from analysis. As such, 418,917 products remained, with 100 unique brand names and 11 distinct countries of origin. Nine products came from developed countries (United States, Japan, Switzerland, Korea, Sweden, Germany, The Netherlands, Italy, and Denmark) and the remaining two from developing nations (China and India). Unfortunately, given that most companies selling on Amazon.com appear to originate in developed nations, the distribution between developed and developing nations was not equivalent. Most products came from the United States (211,109 items), followed by Japan (95,337 items), Switzerland (35,833 items), Korea (18,816 items) and China (30,833 items). The remaining countries had roughly 10,000 products or less per country, with India having significantly fewer products (623 items). In addition, the data were further reduced to ensure that each country corresponded to at least two or more distinct brands. Countries with fewer than two discrete brand names were dropped from analysis because using one brand to represent a country runs the risk of simply cataloguing how well one particular brand performed from a given country. After dropping countries with fewer than two unique brands, only six countries remained: The United States (211,109 items), Japan (95,337 items), China (30,833 items), Korea (18,816 items), Germany (7,566 items), and The Netherlands (5,516 items). In the final dataset, there are 369,177 distinct items, with 338,344 products corresponding to developed nations and 30,833 products reflecting developing nations.

Fashion

For the fashion category, there were 49,399 items and a total of 1,182 unique brand names. Again, to maintain a homogeneous sample and ensure that all brands had more than one instance, only brands occurring over 50 times were included into the sample corpus. This reduced the number of brand names to 183. Additionally, to guarantee that each brand had multiple products, only companies that offered more than five products were included in the data, resulting in 121 unique fashion brands. Finally, the country of origin for each of these unique fashion brand names was manually coded. Like the electronics data, some of the brands extracted from data cleansing had no discernable country-of-origin and, thus, were dropped from the set. After manually coding in all countries of origin for the remaining brands, a total of 62 unique instances were left.

These brands along with their corresponding countries of origin were then appended back into the original fashion dataset. All items that were not included in the appended brand list were subsequently dropped from analysis. As such, 16,568 fashion products remained, with 62 unique brand names. There were nine distinct countries of origin: six from developed countries (United States, Japan, Italy, Australia, Germany, and Belgium) and the remaining three from developing nations (China, Malaysia, and Egypt). Unfortunately, given that most companies selling on Amazon.com appear to originate in developed nations, the distribution between developed and developing nations was not equivalent. The majority of products came from the United States (12,148 items), followed by Japan (3,425 items), Italy (427 items), China (179 items), and Germany (107 items). The remaining countries had roughly 100 products or less per country, with Egypt having significantly fewer products (16 items) for fashion items. Again, in the final step of the data cleansing process, countries with fewer than two distinct brand names were dropped from analysis because using one brand to represent a country runs the risk of simply cataloguing how well one particular brand performed from a given country. After dropping countries with fewer than two unique brands, only four countries remained: US (12,148 items), Japan (3,425 items), Italy (427 items), and China (179 items). In the final dataset, there are 16,179 distinct data points, with 16,000 products corresponding to developed nations and 179 products reflecting developing nations.

Study 2 – The Effect of Brand Equity on Amazon Ratings

Design

Like study 1, study 2 follows a natural experiment pattern of consumer behavior on Amazon. The same data from study 1 were once again used to investigate how brand names may influence the overall rating of Amazon products. Specifically, we explore how higher brand equity may result in different average ratings for products with low brand equity. Again, as Amazon customers often leave reviews of their purchased on a scale from 1 star (low) to 5 stars (high), Amazon provides a unique natural experiment setting to examine the effects of brand name on customer perceptions of the product or brand. Considering these conditions, study 2 follows a 2 by 2 (brand equity: high vs. low x consumer rating: low vs. high) natural design. We theorize a more notable brand name will result in different average ratings than generic brands; explicitly, we hypothesize that a more noteworthy brand will be associated with higher average reviews on Amazon compared with a less familiar brand.

Participants

Participants were not actively recruited for this research as it follows a natural experiment design. Instead, participants consisted of Amazon customers who had purchased a product from Amazon.com and left a product rating on a scale of 1 to 5. In total, we analyzed data from 385,356 Amazon consumers (combined electronics and fashion data).

Measures

Economic Development

In this study, we explore brand equity (e.g., a well-known brand versus a relatively unknown brand) as an independent variable. Particularly, we explore how a brand name may be predictive of consumer ratings on an e-commerce site. To separate highly familiar brands from generic brands, two independent coders coded a list of brand names from the electronics and fashion datasets, respectively. Intercoder reliability scores were calculated for the independent classifications. When a satisfactory score was obtained, these classifications were implemented for exploring how brand familiarity affects customer ratings.

Consumer Rating

A natural real-world method is employed to measure the dependent variable, customer perceptions of the product or brand. Because consumers are asked to leave a rating out of 5 stars on Amazon, with 1 star reflecting a low rating and 5 stars reflecting a high rating, we are easily able to gauge whether consumers held positive, negative, or neutral, perceptions of their purchased item.

Materials

Again, because study 2 follows a natural experiment design, we use materials archived online. We retrieve Amazon data on the items that consumers purchased, the date and time of purchase, the price, and the brand (as the indicator of country-of-origin). Additionally, we only use datasets containing ratings for the products so we are naturally able to obtain ratings through the Amazon data archive. Additionally, we utilize the list of independently coded name brands (classified as familiar or non-familiar).

Procedure

Study 2 uses the cleaned datasets resulting from the data reduction phase of study 1. However, because the objective of study 2 is to further probe how brand name influences the score a customer leaves on Amazon, more data manipulation had to be conducted for study 2 in order to produce relevant results. First, brand names were randomly selected from each of the countries in the fashion and electronics datasets, respectively. Following this, two independent coders attributed a label of 0 (not familiar), 1 (familiar), or 2 (unsure) for brands in both categories. The labels were scored for agreement using Cohen's kappa (Cohen, 1960) and when percent agreement reached a minimum of .80 (substantial), the resulting classifications were implemented in the analysis. Using the cleaned fashion data from study 1, study 2 began with a set of 369,177 products for the electronics category. These data reflected electronics products that met the following criteria: 1) brand occurred a minimum of 500 times; 2) each brand offered at least 20 products; 3) brands had a clear country-of-origin; and 4) countries correspond to at least two brand names. The resulting electronics set contained the following six countries: United States (211,109 items), Japan (95,337 items), China (30,833 items), Korea (18,816 items), Germany (7,566 items), and The Netherlands (5,516 items). In the final electronics set from study 1, there are 369,177 distinct items, with 338,344 products corresponding to developed nations and 30,833 products reflecting developing nations. Originally, the number of unique electronics brand names amounted to 99 but this was reduced to 23 brands to create more evenly distributed groups. Brand names were randomly selected from each of the six countries where there was a maximum cap of 10 brands per country, which helped prevent data from being too skewed by disproportionally high numbers of brands from one country.

All the brand names collected from this process were independently coded as 0 (not familiar), 1 (familiar), or 2 (unsure) as a measure of brand equity. After coding the remaining 23 brands into familiar or unfamiliar categories, a Cohen's kappa (Cohen, 1960) was calculated to represent levels of agreement for brand familiarity. There was substantial agreement between the two coders' judgments: $\kappa = 1.0$ (95% CI, 1.00 to 1.00), $p < .0005$. All brands matching the names on this list were kept in the data and classified according to the coders' classification of familiar or unfamiliar. Through this categorization process, 26 brand names were used in the analysis, with 13 brands representing familiar brands and the remaining 13 representing unfamiliar ones.

Fashion

Using the cleaned fashion data from study 1, study 2 began with a set of 16,179 products for the fashion category. These data reflected fashion products that met the following criteria: 1) brand occurred a minimum of 50 times; 2) each brand offered at least five products; 3) brands had a clear country-of-origin; and 4) countries correspond to at least two brand names. The resulting set contained four countries: US (12,148 items), Japan (3,425 items), Italy (427 items), and China (179 items). In the fashion dataset from study 1 there are 16,179 distinct data points, with 16,000 products corresponding to developed nations and 179 products reflecting developing nations. Originally, the number of unique fashion brand names amounted to 57, but this was reduced to 22 brands to create more evenly distributed groups. Brand names were randomly selected from each of the four countries, where there was a maximum cap of 10 brands per country, which helped prevent data from being too skewed by disproportionally high numbers of brands from one country.

All the brand names collected from this process were independently coded as 0 (not familiar), 1 (familiar), or 2 (unsure) as a measure of brand equity. After coding the remaining 22 brands into familiar and unfamiliar categories, a Cohen's kappa (Cohen, 1960) was calculated to represent levels of agreement for brand familiarity. There was substantial agreement between the two coders' judgments: $\kappa = .818$ (95% CI, .602 to 1.00), $p < .0005$. After discussing the errors, the coders recoded the content until perfect agreement was achieved. All brands matching the names on this list were kept in the data and classified according to the coders' classification of familiar or unfamiliar. Through this categorization process, 22 brand names were used in the analysis, with 11 brands representing familiar brands and the remaining 11 representing unfamiliar ones.

Study 3 – Development Levels on Ratings over Time

Design

Study 3 follows a nonexperimental design as only correlational time series data were analyzed. In this study, we examine the relationship between two variables of interest, ratings and time for developing nations, without controlling for extraneous variables. In particular, this is an observational analysis of how Amazon ratings for products originating from developing nations have changed over time.

Participants

Participants were not actively recruited for this study as it follows an observational nonexperimental design. Instead, participants consisted of Amazon customers who had purchased a product from Amazon and left a rating on a scale of 1 to 5. In total, we analyzed data from 30,833 Amazon consumers (including only developing nations for electronics data).

Measures

Time

In this study, we explore how brands from developing nations may change over time on Amazon. Interestingly, most of the developing nation brands on Amazon originate in China, so we will primarily explore how time is related to different ratings on Amazon for Chinese products.

Consumer Rating

A natural real-world method is employed to measure the dependent variable, customer perceptions of the product or brand. As consumers are asked to leave a rating out of 5 stars on Amazon, with 1 star reflecting a low rating and 5 stars reflecting a high rating, we are easily able to gauge whether consumers held positive, negative, or neutral perceptions of their purchased item.

Materials

As in previous studies, we once again utilize materials archived online for study 3. For this particular study, we use only the cleaned electronics dataset from the previous study as the fashion set contained only 179 elements from developing countries (compared with 30,833 in the electronics data).

Procedure

As the fashion dataset was highly skewed, containing only 179 developing nations compared with over 16,000 elements for developing nations, the fashion dataset was dropped in study 3 as it was highly biased by the unequal sample sizes. Instead, only the electronics dataset, which contains over 369,177 distinct items, with 338,344 products corresponding to developed nations and 30,833 products reflecting developing nations, was used in study 3. This larger sample and distribution of the data provides a more robust body of information on how development levels may influence ratings over time. The cleaned data from the electronics set were separated based on country of origin, with developed nations (e.g., US, Japan, Korea, Germany, and The Netherlands) in one set and developing nations (e.g., China) in another. Because hypothesis 3 posits that brands from developing nations will see a steady increase in ratings over time, only the developing nations data containing over 30,000 instances were used to see how scores on Amazon change over time.

RESULTS

Descriptive Statistics

Electronics

Prior to running correlation tests, descriptive statistics representing the average rating received on Amazon for each country (mean overall ratings per country) were separately calculated. For brands in the electronics category, the range of average Amazon ratings grouped by country ranged from 4.02 stars to 4.37 stars on a 5-point scale. Specifically, Chinese brands received mean overall of 4.14 stars (SD = 1.24 stars) and the average cost per item was roughly 134.06 dollars. German brands earned an average rating of 4.13 (SD = 1.16) and the typical cost for German items was only 86.28 dollars. Japanese brands received the highest mean overall rating with an average score of 4.37 (SD = 1.04), but Japan had the highest mean cost per product at 197.65 dollars. Korea had the second highest overall average rating with a mean of 4.35 stars (SD = 1.10) and the second highest average cost per product at 168.85 dollars. The Netherlands, surprisingly, earned the lowest overall mean rating, with only 4.02 stars (SD = 1.24), but also offered the cheapest products, on average, costing roughly 57.93 dollars. Finally, products from the US received a mean overall score of 4.23 (SD = 1.18) with the items costing, on average, 73.45 dollars (see Table 1). In electronics, brands from The Netherlands had the lowest average scores (mean = 4.02, SD = 1.24) while brands from Japan reflected the highest overall ratings (mean = 4.37, SD = 1.04).

TABLE 1
AVERAGE RATING BY COUNTRY (ELECTRONICS)

	Amazon Rating					
Country of Origin	Count	Mean	Std	Min	Max	Price*
China	30833	4.14	1.24	1.00	5.00	134.06
United States	211109	4.26	1.18	1.00	5.00	73.44
Japan	95337	4.37	1.04	1.00	5.00	197.65
Germany	7566	4.13	1.16	1.00	5.00	86.28
Korea	18816	4.35	1.10	1.00	5.00	168.86
Netherlands	5516	4.02	1.24	1.00	5.00	57.93
Average	61529.50	4.21	1.18	1.00	5.00	10

**Price in dollars*

Fashion

For brands in the fashion category, similar trends to the electronics brands regarding average ratings grouped by country of origin were found. It was observed that most countries received similar reviews on Amazon, with no distinct outliers for fashion brands. On a 5-point scale, Chinese products earned a mean overall rating of 4.43 with an average cost of 42.25 dollars per item, Italy received a mean overall rating of 4.28 with an average cost of 43.45 dollars per item, Japanese items had an average rating of 4.43 with items costing roughly 69.38 dollars, and finally, products from the US had a mean overall rating of 4.34 and products generally cost 33.27 dollars (see Table 2). Each country had a minimum score of 1 and a maximum score of 5 as no null values were included for analysis. Italy received the lowest average rating (mean = 4.23, SD = 1.09), whereas China (mean = 4.43, SD = 0.95) and Japan (mean = 4.43, SD = 1.09) were tied for highest overall ratings.

TABLE 2
AVERAGE RATING BY COUNTRY (FASHION)

	Amazon Rating					
Country of Origin	Count	Mean	Std	Min	Max	Price*
China	179	4.43	0.95	1.00	5.00	42.25
United States	12148	4.34	1.04	1.00	5.00	33.27
Japan	3425	4.43	0.93	1.00	5.00	69.38
Italy	427	4.23	1.09	1.00	5.00	43.45
Average	4044.75	4.36	1.00	1.00	5.00	119.70

**Price in dollars*

Country Origin on Ratings

Because the datasets were rather large, data had to be split based on category to minimize potential false positives. To test hypothesis 1 (Amazon consumers will rate products from developed (developing) countries more (less) favorably than products from developing (developed) countries) for electronics and fashion categories, products were first labeled according to country-of-origin and subsequently coded as either hailing from a developing or a developed nation. After dividing the products and reviews based on country-of-origin development level, the data were subject to a one-way analysis of variance (ANOVA) to explore whether the difference between overall ratings consumers attributed to brands from developed nations differed significantly from developing nations' products. Partial eta-squared analyses were also performed on the data to provide an indication of the effect size of the ANOVA results.

For electronics products, the results suggest that there was a significant difference between average overall ratings for developed nation brands versus developing nation brands as determined by a one-way ANOVA ($F(1, 369175) = 465.99$, $p = 2.76e^{-103}$, $\eta p^2 = 0.00126$). However, given the magnitude of the data (i.e., 369,177 distinct data points) and the small *p*-value, an effect size analysis was performed to validate that the country-of-origin effect was not simply due to the size of the data, which naturally amplifies even miniscule differences between two groups. As such, the data were further subject to a partial eta-squared analysis to calculate the effect size of the data. The result of the partial eta-squared indicates that even though the ANOVA suggests the overall ratings from developed and developing nation brands are statistically significant, the difference is trivial because anything less than a 0.01 effect size is considered to be a trivial effect, and the result of the partial eta-squared ($\eta p^2 = 0.00126$) suggests the quantity of data augmented inconsequential differences between the means. Hypothesis 1 for the electronics data was not supported; that is, a product's country-of-origin (developed vs. developing) does not provide more than a trivial effect for differences in consumer ratings on Amazon.

Further, to verify that the country-of-origin effect would not be significant given a smaller dataset for electronics, a random sample of 1,000 points was extracted from the original corpus. These randomly selected values were subject to an ANOVA to validate the assumption that the large amount of data for electronics may have magnified a trivial effect, leading to a false positive. The results of the second ANOVA on the electronics data confirmed the original finding that the difference between overall ratings for developed and developing nation brands is not statistically meaningful (F(1, 998) = 0.86, p = 0.35, ηp^2 = 0.0009). Using a smaller subset of the original data, it is evident that the original sample may have been too large, leading to significant differences despite a small effect size. The analysis run on the subset indicates that the effect of country-of-origin is not significant and, if significance were detected, the effect would be trivial (ηp^2 = 0.0009). Again, hypothesis 1 for the random subset of electronics data was not supported; that is, a product's country-of-origin (developed vs. developing) does not provide more than a trivial effect for differences in consumer ratings on Amazon.

After dividing the products based on country-of-origin, the fashion data also underwent a one-way ANOVA to explore whether the average overall ratings customers attributed to Amazon products differed significantly based on country-of-origin. However, as the size of the data was much smaller in comparison with the electronics data (i.e., only 16,179 distinct points), the ANOVA was less sensitive to predicting a significant effect based on trivial differences. For the fashion data, the results of the one-way ANOVA indicate that there was not a significant difference between average overall ratings for developed nation brands versus developing nation brands (F(1, 16177) = 0.84, p = 0.36, $\eta p^2 = 5.16e^{-05}$). Although the analysis of variance did not prove statistically significant, a partial eta-squared was still performed on the data to measure effect size. Again, it was found that any differences between overall ratings for developed nations and developing nations would be considered less than trivial ($\eta p^2 = 5.16e^{-05}$). Taken together, these results do not support hypothesis 1 for the fashion data. There is no significant relationship between a product's country-of-origin and its propensity to receive positive or negative reviews on Amazon (see Figure 1 in Appendix).

TABLE 3
ONE-WAY ANOVA ON COUNTRY-OF-ORIGIN WTH AMAZON RATINGS

	One-way Analysis of Variance (ANOVA)				
	SS	df	F	PR(<P)	η_p^2
COO ($Electronics^1$)*	35.624	1.0	465.997	$2.76e^{-103}$***	0.0013
Residuals	28222.259	369175			
COO ($Electronics^2$)*	0.076	1.0	0.865	0.353	0.001
Residuals	88.319	998			
COO ($Fashion$)	0.859	1.0	0.835	0.361	$5.16e^{-05}$
Residuals	177.010	16177			

**Subscript 1 denotes original electronics data, subscript 2 denotes subset of 1,000*
****Significant at p < .001*

Brand Name on Ratings

To test hypothesis 2, the ratings for products associated with name brand products were compared against the ratings for products associated with unknown brand names. Using the reduced data from study 1 and further reducing the dataset to only 26 brand names (13 familiar, 13 unfamiliar), the electronics brands were subject to a one-way ANOVA. The results of the one-way ANOVA indicate a significant difference between average overall ratings for highly familiar brands ($M_{familiar}$ = 4.32, $M_{unfamiliar}$ = 4.17) versus unfamiliar brands (F(1, 130089) = 247.552, $p = 9.99e^{-56}$) in the full electronics dataset. To ensure that the massive size of the data did not cause the significant result, the same data were reduced to a random sample of 5,000 products. The result of the one-way ANOVA on the reduced electronics data confirms that the difference between familiar and unfamiliar brands is significant (F(1, 4998) = 14.287, p = 0.0002). Taken together, these findings suggest that the overall ratings attributed to well-known brands are higher than those attributed to more generic brands. Hypothesis 2 is confirmed for products in the electronics category.

The same analyses were further performed on fashion category items. Using the reduced dataset from study 1, the data were further reduced to include only 22 brand names (11 familiar, 11 unfamiliar). These data were subject to a one-way ANOVA to explore whether the overall ratings for name brand items would significantly differ from non-branded items. The result of the one-way ANOVA on the fashion data indicates that there is a statistically significant difference between ratings for familiar brands ($M_{familiar}$ = 4.45, $M_{unfamiliar}$ = 4.29) versus unfamiliar brands ($F(1, 7360) = 47.993$, $p = 4.78e^{-12}$). Once again, to ensure that the large size of the data did not produce false positives, the ANOVA was run on a random subset of 5,000 items from the fashion category. The resulting analysis indicates that there is a significant difference for the subsample data ($F(1, 4998) = 11.308$, $p = 0.0007$). These findings suggest that overall ratings attributed to known brands are significantly higher than those characterized as unfamiliar brands, confirming hypothesis 2 for items in the fashion category (see Figure 2 in Appendix).

TABLE 4
ONE-WAY ANOVA ON BRAND FAMILIARITY WTH AMAZON RATINGS

	One-way Analysis of Variance (ANOVA)			
	SS	df	F	PR(<P)
Brand ($Electronics^1$)*	300.123	1.0	247.552	$9.99e^{-56}$ ***
Residuals	157715.666	130089		
Brand ($Electronics^2$)*	18.187	1.0	14.287	0.0002 ***
Residuals	6362.392	4998		
Brand ($Fashion^1$)*	45.895	1.0	47.933	$4.78e^{-12}$ ***
Residuals	7047.048	7360		
Brand ($Fashion^2$)*	1.251	1.0	11.308	0.0007 ***
Residuals	553.104	4998		

**Subscript 1 denotes original dataset, subscript 2 denotes subset of 500*
****Significant at p < .001*

Development Level on Ratings over Time

In testing hypothesis 3 (ratings for products in developed countries would increase over time), an ordinary least squares regression was run on the cleaned data from study 1. Because fashion brands from study 1 did not contain a large enough sample of items from developing nations (i.e., 179 items from China), and over 40,000 items data points were used in the electronics data while testing hypothesis 3. The data were subject to a multiple linear regression (Cronk, 2012) where overall ratings reflected the outcome variable and year along with price reflected the independent variables. The results of the multiple linear regression calculated to predict ratings based on the year of purchase and price of the product reveal a significant equation ($F(2, 30985) = 57.99$, $p < 0.000$) with an R^2 of 0.004. The predicted average Amazon rating is equal to −98.408 + 0.0509 (year) − $1.265e^{-05}$ (price), where year is measured in years and price is measured in dollars. The average overall rating increased 0.0509 stars for each year over time and decreased $1.265e^{-05}$ for every dollar in price. Both price and year were significant predictors of overall ratings on Amazon.

These findings suggest that overall ratings for developing nation brands do increase with an increase in year, confirming hypothesis 3. As hypothesized, it appears developing nation brands, mostly represented by brands hailing from China in this study are quickly gaining on developing nation brands in favorability over time. The overall average ratings for brands coming from developed countries increase, albeit slightly, for an increase in year. Additionally, we find that the rate of increase for average ratings over time appears slightly faster for developing nation brands compared with developed nation brands, suggesting that brands from developing countries, such as China, may soon reach similar ratings to developed nations on Amazon's e-commerce shopping marketplace (see Figure 3 in Appendix).

TABLE 5
OLS REGRESSION ON AMAZON RATINGS BY YEAR FOR DEVELOPED COUNTRIES

Dep. Variable:	overall	**R-squared:**	0.004
Model:	OLS	**Adj. R-squared:**	0.004
Method:	Least Squares	**F-statistic:**	57.99
Date:	Fri, 03 Aug 2018	**Prob (F-statistic):**	7.25e-26
Time:	13:44:01	**Log-Likelihood:**	-50642.
No. Observations:	30988	**AIC:**	1.013e+05
Df Residuals:	30985	**BIC:**	1.013e+05
Df Model:	2		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
Intercept	-98.4083	9.679	-10.167	0.000	-117.379	-79.437
year	0.0509	0.005	10.593	0.000	0.042	0.060
price	-1.265e-05	5.46e-05	-0.232	0.817	-0.000	9.44e-05

Omnibus:	6240.022	**Durbin-Watson:**	1.826
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	10709.765
Skew:	-1.386	**Prob(JB):**	0.00
Kurtosis:	3.784	**Cond. No.**	2.77e+06

DISCUSSION

This research offers notable contributions to both scholarship in international marketing and advertising as well as to business practitioners. The results from this study are particularly relevant as consumerism is moving rapidly toward digitization and e-commerce. Although extant literature on the country-of-origin effect explored how a product's country alliance, specifically the economic development level of that country, affects consumer evaluations (Hu & Wang, 2010; Andéhn et al., 2016), few studies investigated the applicability of the COO effect to online shopping and retail. As such, we address this gap in the literature by replicating the current body of research on how level of economic development (developed vs. developing) of a country may impact consumer product evaluations on the online shopping platform, Amazon.com.

Using the country-of-origin effect as our theoretical backdrop, our study reveals that consumers significantly attribute higher scores to products with a developed country-of-origin rather than developing country-of-origin. However, although Amazon customers affixed higher ratings to brands and items from developed nations, the magnitude of the effect was less than trivial, suggesting that although the two groups differ, country-of-origin may be less important on an aggregate e-commerce site. Our second study reveals that consumers significantly attach higher ratings to well-known brands compared with lesser known brands on Amazon, signifying perhaps that country-of-origin falls second to brand notoriety in online shopping. For instance, several brands of Chinese origin (e.g., D-Link, Acer, and Asus) were well received by Amazon shoppers, with many giving these brands high reviews despite their country-of-origin. Further, we find that brands from developing nations are generally receiving higher ratings over time; although the increase over time is not momentous, it is significant.

Taken together, the results of our studies indicate that the e-commerce environment differs from traditional brick-and-mortar settings. Applying theories primarily constructed for traditional retail, even those vetted and replicated, to e-commerce markets may lead to potential inconsistencies between theory and practice; customers, for instance, do not always behave the same in real-life as they do online. As such, our research highlights the necessity of scholarly replications of business research to the increasingly mainstream e-commerce environment. Our findings advance the body of knowledge on the country-of-origin theory as we explore how the theory replicates on Amazon.com. We find that country-of-origin is not necessarily that predictive of ratings as customers only attribute slightly higher scores to brands hailing from developed nations. We also find brand familiarity to be an integral component on Amazon as consumers tend to provide higher ratings to familiar brands, regardless of the brand's origin. Past researchers have also found brand familiarity to be a relevant cue when considering extrinsic cues such as country-of-origin (Jo, Nakamoto, & Nelson, 2003; Koschate-Fischer et al., 2012); however, previous studies have yet to analyze the effect in e-commerce.

Although our findings in this research offer valuable insight to both scholars and practitioners alike, as with all research, it is not without its limitations. In our study, we used a natural experiment setting to test our hypotheses based on real consumer data from Amazon. As such, we had less control over the other variables surrounding purchase and ratings. This natural experiment, for instance, resulted in skewed data where more information for developed countries was accessible compared with developing countries. Further research should aim to replicate our results using a controlled environment to explore whether similar conclusions surface. Additionally, we only explore two product categories in our study, fashion data and electronics data. Amazon.com offers products ranging from household goods to fresh produce, but our research was limited to only two narrow product categories. Future research should investigate how well our findings hold across a wider range of product categories. Finally, our study only investigated COO and familiarity effects on Amazon.com although several other e-commerce sites exist. To explore how well these effects replicate on different platforms and across cultures, future research should explore data from other e-commerce sites such as Alibaba or Taobao.

REFERENCES

Andéhn, M., Nordin, F., & Nilsson, M.E. (2016). Facets of country image and brand equity: Revisiting the role of product categories in country-of-origin effect research. *Journal of Consumer Behavior*, 13(3), 225-238.

Bristow, D.N., Schneider, K.C., & Schuler, D.K. (2002). The brand dependence scale: Measuring consumers' use of brand name to differentiate among product alternatives. *Journal of Product & Brand Management*, 11(6), 343-356.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46

Cronk, B.C. (2012). *How to use SPSS Statistics: A Step-by-step Guide to Analysis and Interpretation*. Pyrczak Pub.

Hu, Y., & Wang, X. (2010), Country-of-Origin Premiums for Retailers in International Trade: Evidence from eBay's International markets. *Journal of Retailing*, 86(2), 200-207.

Jo, M.S., Nakamoto, K., & Nelson, J.E. (2003). The Shielding Effects of Brand Image against Lower Quality Countries-of-Origin in Global Manufacturing. *Journal of Business Research*, 56(8), 637–46.

Koschate-Fischer, N., Diamantopoulos, A., & Oldenkotte, K. (2012). Are Consumers Really Willing to Pay More for a Favorable Country Image? A Study of Country-of-Origin Effects on Willingness to Pay. *Journal of International Marketing,* 20(1), 19-41.

Kotler, P., & Gertner, D. (2002). Country as brand, product, and beyond: a place marketing and brand management perspective. *Brand Management,* 9(4-5), 249-261.

Koubaa, Y., Methamem, R.B., & Fatiha, F. (2015). Multidimensional structures of brand and country images, and their effects on product evaluation. *International Journal of Market Research,* 57(1), 95-124.

Kristensen, T., Gabrielsen, G., & Jaffee, E.D. (2014). Is familiarity a moderator of brand/country alliances in consumer choice? *Transnational Marketing Journal*, 2(2), 61-77.

Lau-Gesk, L., Meyers-Levy, J., & Deighton, J. (2009). *Journal of Consumer Research*, 36(4). 585-599.
Magnusson, P., Westjohn, S.A., & Zdravkovic, S. (2011). What? I thought Samsung was Japanese: accurate or not, perceived country of origin matters. *International Marketing Review*, 28(5), 454–472.
Martin, B.A.S., Lee, M.S.W, & Charlotte, L. (2011). Countering negative country of origin effects using imagery processing. *Journal of Consumer Behavior*, 10(2), 80-92.
O'Shaughnessy, J. & O'Shaughnessy, N.J. (2000). Treating the nation as a brand: some neglected issues. *Journal of Macromarketing*, 20(5-6), 56–64.
Rhee, S.E., & Jung, W.S. (2018). Brand familiarity as a moderating factor in the ad and brand attitude relationship and advertising appeals. *Journal of Marketing Communications*, 1- 15.
Tan, C.T. & Farley, J.U. (1987). The impact of cultural patterns on cognition and intention. *Journal of Consumer Marketing*, 13(3), 27–42.
Usunier J.C. (2011). The shift from manufacturing to brand origin: suggestions for improving COO relevance. *International Marketing Review*, 28(5), 486–496.
Verlegh, P., & Steenkamp, J. (1999). A review and meta-analysis of country of origin research. *Journal of Economic Psychology*, 20(5), 521–546.
Visbal, O., Herrera-Mendoza, K., Orozco-Acosta, E., & Herzberg, M. (2017). The influence of Country of Origin Cues on Product Evaluation: Evidence from Swiss and German Consumers. Journal of Technology Management & Innovation, 12(2), 18-24.
Yan, Y., Zhao, R., & Liu., Z. (2018). Strategic introduction of the marketplace channel under spillovers from online to offline sales. *European Journal of Operational Research*, 267(1), 65-77.

APPENDIX

FIGURE 1
AVERAGE RATINGS BY DEVELOPMENT LEVEL

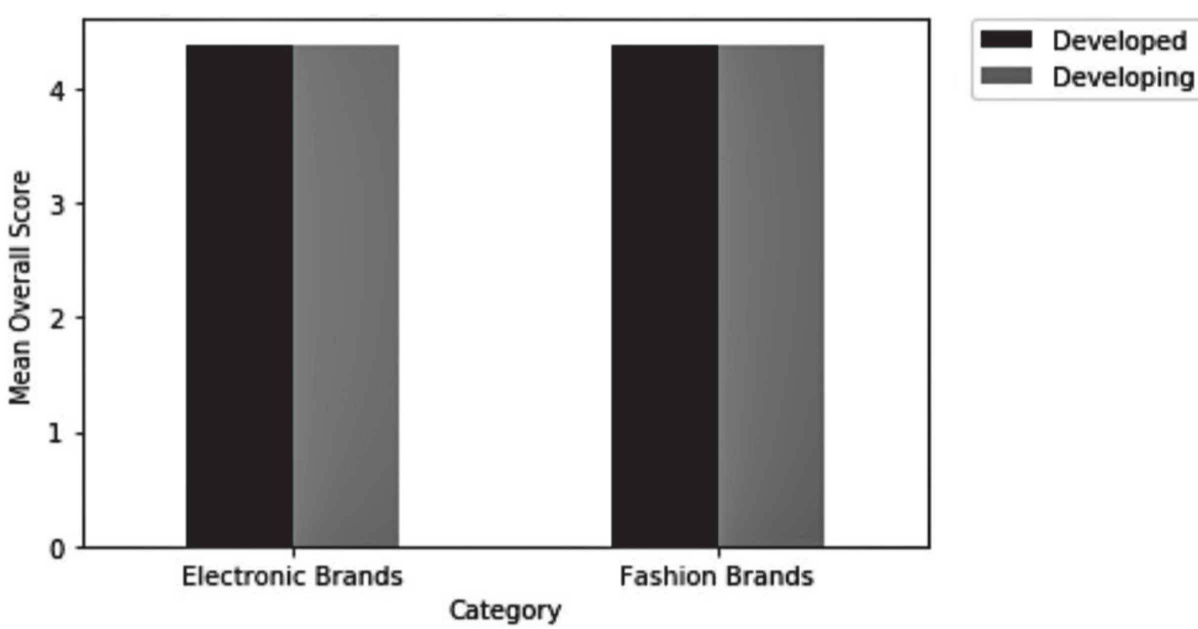

FIGURE 2
AVERAGE RATINGS BY BRAND FAMILIARITY

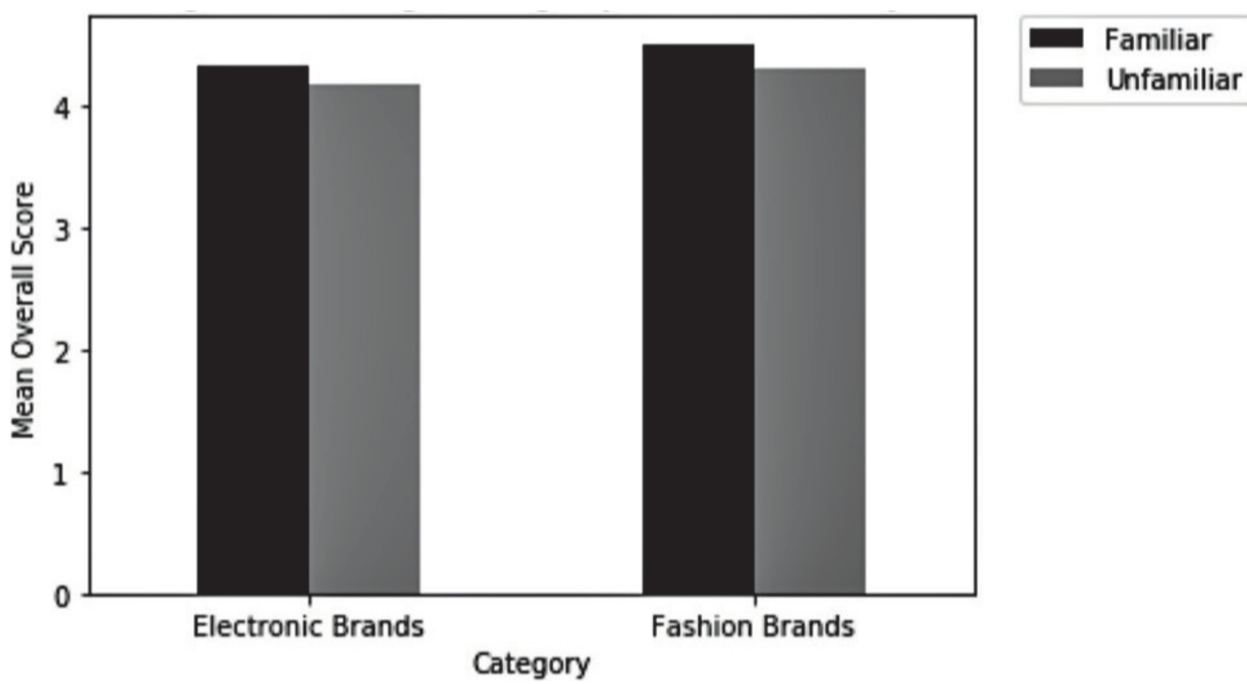

FIGURE 3
AVERAGE RATINGS BY YEAR BASED ON DEVELOPMENT