Submitted to *Management Science*
manuscript

Do Noisy Customer Reviews Discourage Platform Sellers? Empirical and Textual Analysis of an Online Solar Marketplace Using Deep Learning

(Authors' names blinded for peer review)

Customer reviews are essential to online marketplaces. However, reviews typically vary. In many online service marketplaces, sellers are active, i.e., they need to make a proposal to serve each customer. In such marketplaces, it is not clear how (or if) the dispersion in customer reviews affects seller's activity level and number of matches in the marketplace. Our paper empirically examines this by considering both ratings and text reviews. To our knowledge, this is the first paper that empirically studies how the review dispersion affects a seller's activity level and the number of matches in an online marketplace with active sellers. Our paper distinctively investigates the relationship between the review dispersion and *supply-side activities* in an online service marketplace. For our study, we collaborated with one of the largest online solar marketplaces in the U.S. that connects potential solar panel adopters with installers. We complement this with public data sets. Our analysis uses the state-of-the-art deep learning based natural-language-processing model BERT developed by Google AI, an advanced clustering algorithm and econometrics methods. We find that the dispersion in customer reviews has a significant and inverted U-shaped effect on an installer's activity level in the online marketplace. Specifically, installer's activity level increases with the review dispersion if and only if the dispersion is below a certain threshold. Furthermore, we identify a significant and inverted U-shaped relationship between the market-level review dispersion and transactions. Our findings provide valuable insights to marketplace operators about the implications of review dispersion for marketplace operations.

Key words: online marketplace, deep learning, natural language processing, clustering, customer reviews, solar PV installation

---

1. Introduction

Online marketplaces are reshaping numerous sectors, ranging from retail to clean technology. In 2019, gross merchandise sales of top online marketplaces across the globe exceeded the astonishing $2 trillion milestone, with a 22% growth (Ali 2020). Online customer reviews, which are evaluations of a product or service by former users, are vital for online marketplaces because they are essential in customer shopping experience. According to the Spielger Research Center (2017), nearly 95% of customers read online reviews before making a purchase. Customers pay attention to online reviews, and reviews can significantly influence customer perception (Askalidis and Malthouse 2016, Park et al. 2007). However, very little is known about the supply-side impacts of online reviews. In many service marketplaces, supply-side participants are *active*. That is, a seller needs to make a proposal to serve each customer. In these marketplaces, it is not clear how (or if) online reviews affect supply-side activities and overall market transactions. Thus, for such marketplaces, understanding the impact of customer reviews is of paramount importance to a marketplace operator. This is the main focus of our work.

Online reviews can be in a rating or text format. Ratings are typically measured on a 1 to 5 scale, with 1 being poor and 5 being excellent, while text reviews include customer sentiments about the product or service in words. There is a growing interest in studying customer ratings in various contexts. The vast majority of this literature investigates how *average* customer ratings impact a *single* firm's sales. Focusing on books and movies, several studies conclude that an improvement in a product's average rating increases its sales (e.g., Chintagunta et al. (2010) and Chevalier and Mayzlin (2006)). Regarding services, Luca (2016) finds that the average rating of a restaurant has a positive impact on its revenue. There are also a few studies that show that the average ratings of a product may not have a significant impact on its sales (e.g., Duan et al. (2008)). In practice, customer ratings typically vary; it is rare to find a product or service whose ratings are all the same. Despite this, surprisingly, the implications of rating dispersion are severely understudied in the literature (see Section 1.2). Our paper contributes to the literature by studying how online review *dispersion* impacts a key behavior of online marketplace sellers and the online marketplace that consists of multiple active sellers. To the best of our knowledge, there is no prior work that investigates this topic.

Our paper considers an online solar marketplace as a context. Solar energy is booming in the world, with a dazzling 34% growth worldwide in 2017 (IEA 2018). In the U.S., the annual generation from solar photovoltaics (PV) increased by nearly a factor of 4 from 2014 to 2019, and is estimated to more than triple from 2019 to 2030 (U.S. EIA 2020b,a). A key contributor to this growth is increasing solar panel adoption by electricity end-users (e.g., residential customers). By adopting solar panels, electricity end-users generate their own electricity, reducing their reliance on electric utility companies. In the U.S., the residential solar capacity increased by a factor of 4.18 from 2014 to 2019 (U.S. EIA 2020c), and is forecasted to grow 25% per year (Weaver 2019, SEIA 2019, Pyper 2018).

Online marketplaces are transforming the rooftop solar panel adoption process across the United States. An online solar marketplace is a digital platform that connects a potential panel adopter with installers, facilitating the adoption process for electricity end-users. Customers are increasingly interested in connecting with rooftop panel installers through online marketplaces. According to a recent report about a leading online solar marketplace, such an interest doubled in 11 major states of the U.S. between 2017 to 2018 (EnergySage 2019). In this paper, we analyze a novel data set we obtained from one of the largest online solar marketplaces in the United States.

The online solar marketplace we study has two salient features. First, for every incoming customer, each installer in a certain region decides whether to serve that customer or not. The installer makes a proposal (bid) if it is willing to serve the customer. Only after the installer's proposal, the installer is listed as available for the customer. This is in contrast to online marketplaces such as Amazon where sellers do not bid for a potential customer. Second, the competition among installers is local. That is, only the installers located in a particular geographical area bid for each customer, and this geographical area is not restricted to city or town boundaries. This is different from online marketplaces like Amazon where the competition among sellers occurs at the entire marketplace level. This difference creates a unique challenge, that is, to identify *local markets* for installers. In our study, we overcome this challenge via an advanced clustering algorithm.

Our paper considers two key metrics: the number of proposals by each installer, which represents the number of customers each installer is willing to serve on the marketplace, and the number of successful proposals - i.e., *matches* - in the marketplace. The former is relevant to the growth prospect of the online marketplace, which is an important measure for investors (Baker 2020). The latter metric matters as it is commonly used in the financial valuation of online marketplaces (Boris 2018, Galston 2017). Hereafter, for brevity, the logged number of proposals by an installer will be referred to as the installer's *activity level*.

Our analysis is centered around the following three main research questions. (i) Does the dispersion in an installer's customer reviews have a significant impact on the installer's activity level in the online marketplace? If so, what is the direction of the impact? (ii) Does the dispersion in competitors' customer reviews have a significant impact on an installer's activity level in the online marketplace? If so, what is the direction of the impact? (iii) How does the review dispersion impact the number of matches in the marketplace? In answering these questions, we consider both ratings and text reviews made by verified buyers. To consider these two formats, in addition to traditional econometrics methods, we employ the BERT technique, which is an advanced natural language processing model implemented by Google in late 2019. To our knowledge, our paper is the first that employs this technique in the OM literature.

1.1. Main Findings and Contributions

Our paper makes four main contributions to the literature. First, to the best of our knowledge, there is no prior work that empirically investigates how the dispersion in customer reviews impacts a firm's activity level (i.e., logged number of proposals) in an online marketplace. Our paper studies this, and shows that the dispersion in an installer's reviews has a significant and inverted U-shaped impact on its activity level in the online marketplace. Thus, a firm's noisy reviews increase the firm's activity level if and only if its review dispersion is lower than a threshold; beyond that threshold, noisy reviews hurt the firm's activity level in the online marketplace.

Second, to our knowledge, our paper is the first that studies how the dispersion in competitor reviews impacts a firm's activity level in an online marketplace. In this context, we find that competitors' rating dispersion has a significant and inverted U-shaped impact on the installer's activity level. This suggests that a firm's and its competitors' rating dispersions have the similar structural impact on the firm's activity level in the online marketplace.

Third, to our knowledge, our paper is the first to empirically analyze how the review dispersion affects the number of matches in an online marketplace where sellers have to make a proposal to win a customer. Regarding this, we identify a significant and inverted U-shaped relationship between the number of matches and the review dispersion at a local market level. This finding has a key implication for an online marketplace operator: Having all sellers with 5 stars might not be favorable to the marketplace operator. Review dispersion up to a particular level can help an online marketplace operator in terms of number of matches.

Fourth, our paper provides a showcase for an advanced clustering method (OPTICS) and a very recent deep-learning based natural-language-processing model (BERT). These methods have not been used in the OM literature yet, and have the potential to facilitate research in various contexts.

1.2. Related Literature

Our paper contributes to the sustainable operations literature by examining an online marketplace that facilitates solar PV adoption. Here, we will only mention the most relevant papers that includes a data analysis. Interested readers can find an excellent review in Lee and Tang (2018). In this stream, various papers analytically study solar and wind technologies while calibrating their models with real-life data (see, e.g., Singh and Scheller-Wolf (2018), Sunar and Swaminathan (2018), Sunar and Birge (2019), and references therein). There are also several papers that empirically study green technologies/solutions. These include carbon abatement solutions (e.g., Blanco et al. (2020), Huang et al. (2020), Blass et al. (2014)), waste exchanges (e.g., Dhanorkar et al. (2015)) and off-grid lighting solutions (e.g., Uppari et al. (2019), Kundu and Ramdas (2019)). To the best of our knowledge, there is no prior work that considers customers reviews and an online solar marketplace in this literature.

Our paper also contributes to the literature on online marketplaces. Moreno and Terwiesch (2014) use a transactional data set from an intermediary for software development services. The authors establish that for a seller, a superior reputation primarily increases its likelihood of winning a business. Bimpikis et al. (2019) use data from a natural experiment in a liquidation auction on a business-to-business platform, and illustrate that the design of the online platform significantly impacts the platform's revenues. Li and Netessine (2020) analyze data from an online peer-to-peer property-rental platform, and show that the market thickness can decrease the number of transactions on the platform. To our knowledge, in this stream, there is no work that studies how review dispersion impacts a firm's activity level (i.e., logged number of proposals) or the number of matches in an online marketplace with active sellers, which are the topics of our study.

There are a few papers that study the impact of rating variability on a firm's sales. However, there is no consensus about the impact. Clemons et al. (2006) find a positive correlation between the rating dispersion and craft beer sales to provide support for a hyper-differentiation marketing strategy in the craft beer industry. In contrast, Zhu and Zhang (2010) show that the rating variation for less popular online games has a negative impact on sales. Luo et al. (2013) find that the dispersion of brand ratings can drastically hurt the firm value while Zhang (2006) concludes that the rating variation does not play a significant role in movie openings. Our paper differs from these studies in several dimensions. First and most importantly, unlike these papers, our paper takes the perspective of a marketplace operator, and studies how review dispersion impacts a key seller action and the number of matches in the marketplace. This is in contrast to the common focus in these papers, which is to understand customer-side impact of online ratings on a single firm. Second, in our setting, the seller must prepare a proposal to win its customers. Such a setting is key to our analysis and not considered by these studies. Third, we consider an online solar marketplace, which differs from studied contexts in essential ways.

Finally, our paper is related to the relative performance feedback (RPF) literature. The vast majority of this literature studies how feedback impacts an *individual worker*'s performance. Performance is context-specific, and hence measured in different ways. For example, in a hospital setting, Song et al. (2017) measure the physician performance by patients' median length of stay, and find that sharing best practices and public RPF (by making physician performance public) improve the performance of low-performing physicians. There are only a few studies that consider firm-level RPF (e.g., Delfgaauw et al. (2013)). Among those, Niewoehner III and Staats (2020) is the only one that considers feedback not tied to financial incentives (e.g., any external prize or penalty as in tournaments). Niewoehner III and Staats (2020) establish that when clinics are informed about their rankings on the flu shot growth, which is the performance metric of interest in their setting, clinics exhibit last-place aversion behavior. Our paper differs from this literature in multiple ways. First, in these studies, providing feedback refers to disclosing a firm's or a worker's relative performance of interest. However, in our study, customer reviews, which are customer feedbacks, do not disclose firms' activity levels and number of matches in an online marketplace, which are the performance metrics of interest. Furthermore, we study the impact of customer review dispersion. To our knowledge, there is no prior work that empirically studies how feedback dispersion affects firms' actions and marketplace operator's performance in this stream.

1.3. Organization of the Paper

The remainder of our paper is organized as follows. Section 2 states our hypotheses and describes several mechanisms by which review dispersion may impact firms' activity levels and the number of matches in the online marketplace. Section 3 explains our data and context, and includes preliminary analysis. Section 4 examines how installer's and competitors' review dispersions impact the installer's activity level in the online marketplace. Section 5 studies the relationship between the market-level review dispersion and the number of matches in the online marketplace. Section 6 employs text-mining techniques to utilize both numerical ratings and text reviews in our analysis. Section EC.2 provides various robustness checks, and shows that our findings are robust.

2. Hypothesis Development

Online reviews play many roles in an online marketplace. One role is related to the brand image of firms. Online reviews have become an integral part of the brand image in the age of e-commerce and digitalization (Chakraborty and Bhat 2018a,b, Gensler et al. 2015). At the same time, online reviews are also seen as a reflection of consumer taste in the market (Clemons et al. 2006). Thus, a firm's review dispersion can have an intricate impact on its marketplace activity level via different mechanisms, which we will explain below.

On one hand, a higher review dispersion of an installer may encourage the installer and thus increase its activity level in the online marketplace. The dispersion in the installer's online reviews can lead to the dispersion in brand image, and that can drastically hurt the firm value (Luo et al. 2013). When faced with a large review dispersion, an installer may be willing to serve more customers to increase the number of its reviews, thereby reducing its review dispersion and improving the consistency of its brand image. As a result, installers with a higher review dispersion may make more proposals to win more customers in the online marketplace.

On the other hand, a higher review dispersion of an installer may also discourage the installer and decrease its activity level in the online marketplace. A higher review dispersion may imply a higher differentiation of customer taste in the market (Clemons et al. 2006). In such a market, making more proposals may impose reputational risks to the installer, potentially due to negative customers or additional polarized reviews. In fact, practitioners warn about these reputational risks, and advise that in the presence of negative customers, efforts to collect more online reviews may results in even more negative reviews for the firm (see, e.g., (Patel 2020)). When faced with reputational risks, firms can be more selective in their project choices to avoid such reputational risks (Demirag et al. 2011, Hirshleifer and Thakor 1992). Thus, as the installer's review dispersion increases, the installer may reduce its activity level to be more selective about which customer to serve in the marketplace.

A higher review dispersion of an installer may also reduce its activity level in the online marketplace through another mechanism. That is, an installer might think that with a higher review dispersion, it is less likely to win a customer compared to its competitors because a higher review dispersion may damage customer perception (Zhu and Zhang 2010). Since making a proposal is costly for an installer, when faced with a lower likelihood of winning customer, the installer may reduce its activity level in the marketplace to avoid expenses that do not generate revenues.

Based on all of these conflicting perspectives, we have the following competing hypotheses:

Hypothesis 1A: *An increase in an installer's review dispersion increases its activity level in the online marketplace.*

Hypothesis 1B: *An increase in an installer's review dispersion decreases its activity level in the online marketplace.*

The dispersion in competitor reviews may also impact an installer's activity level. Competitors, as will be explained in detail, are a set of installers who share the same local market and potential customers. Similar to our earlier discussion, an increase in the dispersion of competitor reviews may be perceived as a signal of a more polarized market. Seeing competitors receive brand-image-damaging reviews can be perceived as a negative about the market. Thus, when competitor reviews become more disperse, to avoid any reputational harm due to polarized reviews or negative customers, the firm may be more conservative in making proposals. That would reduce the installer's activity level in the online marketplace.

At the same time, a higher dispersion in competitor reviews can also increase the installer's activity level in the marketplace through another mechanism. A higher dispersion in competitor reviews might hurt the competitor's brand image, and may negatively impact the customers' perception about the competitors' service (Chakraborty and Bhat 2018a,b, Zhu and Zhang 2010). This may increase the installer's likelihood of winning any customer compared to its competitors if the installer makes a proposal (Demirag et al. 2011, Moreno and Terwiesch 2014). Given the higher likelihood of winning, it may be favorable for the installer to bid for more customers so as to improve its sales. That would increase the installer's activity level in the marketplace. In light of these, we have two competing hypotheses:

Hypothesis 2A: *An increase in competitors' review dispersion increases the installer's activity level in the online marketplace.*

Hypothesis 2B: *An increase in competitors' review dispersion decreases the installer's activity level in the online marketplace.*

Apart from the impact on the installer behavior discussed above, the review dispersion can also influence the number of transactions in the online marketplace. A match, i.e., an agreement between a customer and any installer, occurs only when the customer is willing to accept an available proposal. A key determinant of a customer's willingness to accept any proposal is the number of proposals she receives from installers, which is a direct consequence of installer activity levels. As explained above, installer activity levels can be impacted by review dispersion via various mechanisms.

By the discussions about Hypotheses 1A, 1B, 2A and 2B, it is not clear how the market-level review dispersion impacts the average number of proposals per customer. There are also conflicting perspectives about how the number of available options affects the customer's willingness to accept any option. On one hand, receiving more proposals might overload the customer, and can decrease the customer's motivation to accept any proposal (Scheibehenne et al. 2010, Iyengar and Lepper 2000). On the other hand, having more proposals may also increase the customer's motivation to accept a proposal because in a larger set of options, the customer might be more likely to find a proposal that better matches to her objective (Scheibehenne et al. 2010, Baumol and Ide 1956). Thus, an increase in the average number of proposals per customer may increase or decrease the customer's willingness to accept any proposal, i.e., the likelihood of a match. Combining all, an increase in market-level review dispersion can increase or decrease the (logged) number of matches in the online marketplace. Thus, we have the following two competing hypotheses:

Hypothesis 3A: *An increase in the market-level review dispersion increases the (logged) number of matches in the online marketplace.*

Hypothesis 3B: *An increase in the market-level review dispersion decreases the (logged) number of matches in the online marketplace.*

3. Data and Setting

For our study, we collaborated with one of the largest online solar marketplaces in the U.S., and obtained proprietary marketplace data from the company. We also complement this data set with Tracking The Sun (TTS) data set from the Lawrence Berkeley National Laboratory. TTS is a comprehensive data set on U.S. solar panel installations. Below, we will provide further details about our data and the setting of the online solar marketplace we study.

3.1. Online Solar Marketplace

The solar marketplace we study is an independent shopping website for electricity end-users (e.g., homeowners) who are interested in adopting solar panels. The marketplace operates in 33 states of the U.S., and allows solar panel installers to maintain a profile, receive information and connect with potential customers in their service areas.

The marketplace operates as follows. First, each customer visits the marketplace website and enters her information, such as the location of her property. Next, installers are informed about the customer's arrival along with her information. Each installer only serves to a particular region. If the customer's location falls into an installer's service area, the installer decides whether to make a proposal to the customer or not. After the customer observes installer proposals she receives, there are two possible outcomes: Either the customer agrees to work with an installer, i.e., there is a successful *match*, or the customer gives up the process, i.e., there is no matching. If the customer ends up working with the installer, she can leave a review that contains text and a rating ranging from 1 to 5 stars. The marketplace verifies customers who leave reviews. Hence, reviews are considered as authentic and not manipulated.

As a result, the key decision for each installer in the marketplace is whether to make a proposal for a potential customer or not. In light of this, we study how the dispersion in customer reviews impacts (i) an installer's *activity level* in the marketplace, which is a logarithmic transformation of the number of proposals the installer makes in a month, and (ii) the number of monthly matches in the marketplace. For the reasons explained earlier, both of these are important metrics for the marketplace operator.

To investigate (i) and (ii), we obtained a unique panel data set from the online solar marketplace. Our data set contains the data of the marketplace's all vetted installers across the U.S. and full record of their customer reviews from January 2013 to April 2018. There are 416 installers in the marketplace, and we have each installer's monthly activities, i.e., the number of proposals made and the number of proposals won by each installer in every month, during the aforementioned time frame. Each review has a rating, text content, time stamp, and the installer ID and name with which the review is associated. We also have the location information of each installer, as illustrated in Figure 1. In the marketplace we study, there is no "closing off" or explicit exit behavior as in physical retail stores. If an installer prefers to quit the online marketplace, the installer simply becomes inactive, making no proposals to potential customers. Our analysis accounts for such behavior.

3.2. Defining Local Market

Solar panel installation is a combination of product and service. As part of service, installers typically visit customer site multiple times. Thus, each installer only operates within a certain geographical area, and installers compete "locally." That is, they only compete with installers that are relatively nearby. The caveat is that there are no available data on the installers' service areas. To capture this practical element, we identify "*local markets*" within the marketplace so that only installers in the same local market compete with each other.

To geographically segment the marketplace into local markets, we divide installers into multiple *clusters* and treat each cluster as a separate local market. Boundaries of local markets cannot be simply defined by city, county, or congressional district borders because it is common for installers to

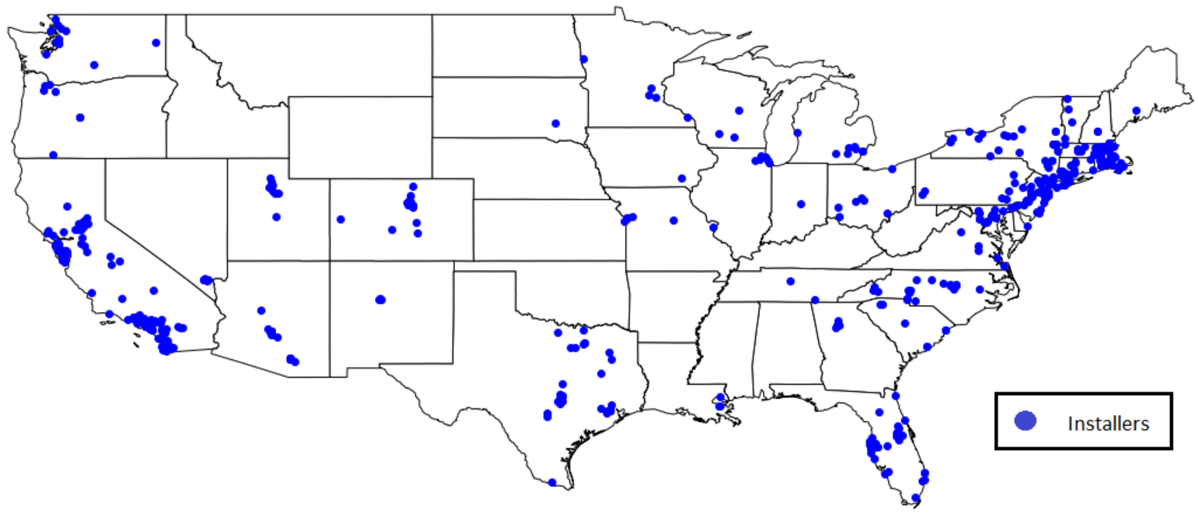

Figure 1 **Installers in our data set**

cross these artificial borders to serve customers. Instead, we use installer locations and an advanced clustering algorithm called OPTICS (*Ordering Points To Identify the Clustering Structure*) to identify local markets.

The OPTICS routine is an unsupervised machine learning algorithm that identifies density-based clusters in spatial data. It is considered to be an extension of various commonly-used advanced clustering algorithms, such as DBSCAN (Kanagala and Krishnaiah 2016). Among others, an important advantage of the OPTICS algorithm is that it does not require fixing the number of clusters before running the algorithm as in k-means clustering method; rather, it identifies the optimal number of clusters using data. Because of this and many other advantages, OPTICS has been applied in various contexts, ranging from political science (Davidson 2019) to geography (Teimouri et al. 2016). To the best of our knowledge, our paper is the first that uses this advanced clustering technique in the OM literature.

In light of these, we create the geographic division of local markets with the following steps. First, we collected the 5-digit zipcode of every installer in the marketplace. Figure 1 displays the location of every installer in our data set. We then converted each zipcode to the representative coordinates based on the data provided by the US Census Bureau (2019). This transformation is necessary to run the OPTICS algorithm on the location data. The OPTICS algorithm uses the maximum distance between two samples in a cluster as an input variable. Based on our conversations with practitioners, we learned that the majority of customers get a quote from an installer within 90 miles distance of their property. Consistent with this, we used 90 miles as the maximum distance input parameter, and the OPTICS algorithm generated 36 clusters. Each of these clusters geographically defines a local market boundary. Figure 2 illustrates the centroid of each of these 36 clusters, which

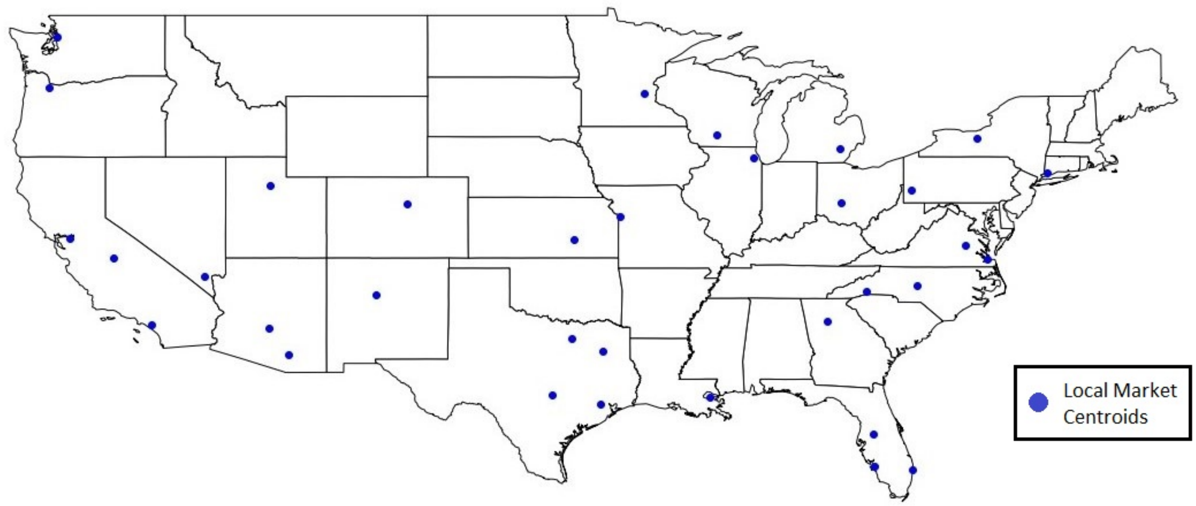

Figure 2 Local Market Centroids

represents the centroid of each local market. Hereafter, for brevity, we refer to local markets simply as "*markets*."

3.3. Measuring Dispersion in Customer Ratings

A key explanatory variable in our regression is the dispersion in customer ratings. This section explains how we measure the rating dispersion. Later, we will also study an extended model by adding the text-based review dispersion as a separate variable in our analysis. Section 6 will explain the state-of-the-art natural language processing model we use to measure the review dispersion based on text data.

We measure the rating dispersion by calculating the *entropy* of ratings. In information theory, the entropy is a common way to measure the information content or dispersion (or uncertainty) in a variable's possible realizations. In our setting, because the marketplace has a 5-star rating system, the entropy of ratings is

$$H(R) = \sum_{j=1}^{5} \text{Prob}(\text{Rating} = j) \log(1/\text{Prob}(\text{Rating} = j)). \tag{1}$$

For example, for a set of 5 reviews each with 4 stars (out of 5 stars), the entropy of ratings $\{4,4,4,4,4\}$ is zero. Alternatively, for a set of 5 reviews with ratings $\{3,5,3,5,4\}$, the entropy of ratings is 1.0549. Although both sets have the same average rating of 4, the latter set of ratings provides more information with a higher dispersion, hence has a higher entropy.

In light of this, we create three variables that measure the rating entropy in different dimensions for each month t. First variable is $\text{Rating\_Entropy\_Self}_{i,t}$, which is the demeaned entropy of ratings associated with installer i up to and including month t. Recalling the market defined in Section 3.2, the second variable is $\text{Rating\_Entropy\_Others}_{i,t}$ that is the demeaned rating entropy of all other installers in installer i's market, up to and and including month t. Our third variable is $\text{Rating\_Entropy\_Mkt}_{m,t}$ that represents the demeaned entropy of all ratings in the market m, up to and including month t. Note that these three variables are centered around their means. This is a standard procedure in settings like ours where the regression includes both linear and quadratic terms of an explanatory variable (see, e.g., Tan and Netessine (2014)). We also checked the robustness of our findings by replacing these variables with their non-demeaned versions in all our econometric analysis, and we find that all of our insights remain the same with non-demeaned variables.

We measure the rating dispersion by calculating entropy rather than variance of ratings. The reason is two folds: First, the entropy measure provides a higher precision for our data than the variance. That is, two installers with very small difference in rating variance tend to show a larger difference in rating entropy. Second, when data display multi-modality as our rating data do, entropy is considered to be a better measure than the variance in capturing the dispersion in data (Smaldino 2013).

4. Installer-Level Analysis & Results

This section examines the following questions: (i) How does the dispersion in an installer's ratings affect its *activity level*, which is the logged number of proposals generated by the installer? (ii) How does the dispersion in competitors' ratings impact the installer's activity level? By studying these questions, we test Hypotheses 1A, 1B, 2A and 2B in Section 2.

We will only use numerical ratings in this section. Later, Section 6 will extend our analysis to include text reviews. Section EC.2 will check the robustness of our findings in various dimensions, and address potential endogeneity concerns in an extended model in Section EC.2.1.

4.1. Regression Equation & Controls

To answer (i) and (ii) above, we consider a regression model where the dependent variable is a natural logarithmic transformation of the number of proposals made by an installer. Formally, indexing installers, months and markets by i, t and m, respectively, the dependent variable in our regression is $\text{Installer\_Activity}_{i,m,t+1}$, which is equal to ln(1 + number of proposals generated by installer i) in the market m during month t+1. We make this transformation because the number of installer proposals has a right-skewed distribution, and log transformation increases the normality of errors, thereby further improves the validity of inference. This is a standard procedure in the literature (see, e.g., Song et al. (2017), Tan and Netessine (2014, 2019), among others). As a robustness check, we also performed the analysis without a logarithmic transformation and found that results are consistent.

Two of our key explanatory variables are $\text{Rating\_Entropy\_Self}_{i,t}$ and $\text{Rating\_Entropy\_Others}_{i,t}$, which are defined in Section 3.3. Because we have competing hypotheses, we also allow for nonlinear relationships between each of these explanatory variables and the dependent variable by including explanatory variables $\text{Rating\_Entropy\_Self}^2_{i,t}$ and $\text{Rating\_Entropy\_Others}^2_{i,t}$ in our regression. Specifically, our regression equation is as follows:

$$\begin{aligned}\text{Installer\_Activity}_{i,m,t+1} = &\beta_0 + \beta_1 \text{Rating\_Entropy\_Self}_{i,t} + \beta_2 \text{Rating\_Entropy\_Self}^2_{i,t} \\ &+ \beta_3 \text{Rating\_Entropy\_Others}_{i,t} + \beta_4 \text{Rating\_Entropy\_Others}^2_{i,t} \\ &+ \text{Controls}_{i,m,t} + \alpha_i + \epsilon_{i,t+1}. \end{aligned} \tag{2}$$

Here, ϵ is the installer-level error term, and represents random factors that are unobservable in the data and affect the installer activity. We run two versions of (2): In one version, we consider α_i as a fixed effect whereas in the alternative version, we consider it as a random effect. To determine which model is more appropriate for our data, we run the Durbin-Wu-Hausman test where the null hypothesis is that the random-effect model is preferred while the alternative is the fixed-effect model. With a p-value < 0.0001, we reject the null hypothesis and conclude that the fixed-effect model is more appropriate. We also establish the significance of the fixed effect in (2) with the F-test ($\chi^2(13) = 44.23, p < 0.0001$). Thus, we focus on (2) with the installer-level fixed effect α_i that controls for time-invariant characteristics of each installer.

In this regression model, the key coefficients of interests are β_1, β_2, β_3 and β_4. The values of these coefficients together with the significance of the associated variables will uncover how the rating entropy impacts the installer's activity level in the online marketplace.

The regression (2) includes various installer-level or market-level control variables ($\text{Controls}_{i,m,t}$). To account for the state-level renewable policy effects, we include state dummies as control variables. We have 33 such variables. We account for the impact of the solar panel prices on installers' activity levels by considering $\text{Price\_Difference}_{i,t}$ as another control variable. In practice, because solar PV systems vary in size, price per KW is a common way to represent the price of the installed solar panel. We collected each installer's price for 1 KW solar panel by matching names and zipcodes and using the TTS data set. Based on this, we compute the variable $\text{Price\_Difference}_{i,t}$ by taking the logarithm of the difference between installer i's price and the average price of its competitors that operate in the same market in month t. We control for the average rating of each installer i as well as the average rating of its competitors in the market for month t by including variables $\text{Average\_Rating\_Self}_{i,t}$ and $\text{Average\_Rating\_Others}_{i,t}$ in (2). We also control for the installer's experience by including the variable $\text{Experience}_{i,t}$ that is the logged number of years the installer has been installing solar systems up to (and including) month t. We collected this information one by one from each installer's website. Another control variable in (2) is $\text{Market\_LogRevenue}_{m,t}$ that measures the logged total dollar value of all solar installations within market m during month t. To create this variable, we augment the market boundaries identified in Section 3.2 with the TTS data. This variable aims to capture total solar installations opportunities in the market, and can be seen as a proxy for the favorableness of the solar installation market. As the final control variable, we consider $\text{Review\_Counts}_{i,t}$ which is the number of each installer i's reviews up to (and including) month t.

Variables	N	Mean	Standard Deviation	Min	Max
Rating_Entropy_Self	4,562	0	0.217	-0.0985	0.9015
Rating_Entropy_Others	4,562	0	0.183	-0.227	0.773
Average_Rating_Self	4,562	4.531	1.316	1	5
Average_Rating_Others	4,562	4.88	0.205	1	5
Review_Count	4,562	5.384	6.836	0	52
Experience	4,562	1.758	0.929	0	3.714
Price_Difference	4,562	-0.0333	0.392	-2.171	3.139
Market_LogRevenue	4,562	12.24	7.9	0	22.3

Table 1 Summary Statistics - Installer Level Analysis

Note: All entropy variables are demeaned.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Rating_Entropy_Self	1							
(2) Rating_Entropy_Others	-0.094	1						
(3) Average_Rating_Self	-0.088	0.069	1					
(4) Average_Rating_Others	0.025	-0.523	-0.017	1				
(5) Review_Count	0.241	0.061	0.205	-0.039	1			
(6) Experience	-0.001	0.143	0.036	-0.062	0.127	1		
(7) Price_Difference	-0.015	-0.043	0.003	0.016	-0.029	-0.027	1	
(8) Market_LogRevenue	-0.026	0.006	-0.029	0.044	-0.086	0.451	-0.059	1

Table 2 Correlation Matrix - Installer Level Analysis

Tables 1 and 2 above present the summary statistics and the correlation matrix. By Table 2, correlations among explanatory variables are relatively low and do not hurt the validity of regression analysis. We also checked VIF scores of variables, and verified that they are all within the suggested range (Hair et al. (2014)), providing further support on the validity of our analysis.

4.2. Results

Columns (I) through (III) of Table 3 present our estimation results for three specifications. Column (III) includes the estimates under the regression model (2), while the estimates in other columns are obtained by considering only some of those explanatory variables in the regression. In particular,

Variables	(I) Installer's Activity Level	(II) Installer's Activity Level	(III) Installer's Activity Level
Rating_Entropy_Self			1.890***
			(0.000)
Rating_Entropy_Self^2			-3.473***
			(0.000)
Rating_Entropy_Others		0.524**	0.488**
		(0.004)	(0.007)
Rating_Entropy_Others^2		-2.533***	-2.593***
		(0.000)	(0.000)
Average_Rating_Self	-0.876***	-0.834***	-0.774**
	(0.000)	(0.000)	(0.001)
Average_Rating_Others	0.000618	0.000236	0.000964
	(0.975)	(0.991)	(0.963)
Review_Count	0.0561***	0.0534***	0.0472***
	(0.000)	(0.000)	(0.000)
Experience	0.218***	0.212***	0.206***
	(0.000)	(0.000)	(0.000)
Price_Difference	0.0593	0.0690	0.0722
	(0.487)	(0.425)	(0.402)
Market_LogRevenue	-0.0168***	-0.0169***	-0.0159***
	(0.000)	(0.000)	(0.001)
Constant	2.690***	2.803***	2.961***
	(0.000)	(0.000)	(0.000)
Observations	4562	4562	4562
Fixed Effect	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes
Adjusted R^2	0.627	0.630	0.633
AIC	13267.2	13234.9	13205.7
BIC	13325.0	13305.6	13289.2

Note: p-value in parentheses; $^{*}p<0.05$;$^{**}p<0.01$;$^{***}p<0.001$

Table 3 Installer Level Analysis

we obtain the estimates in column (I) by excluding all entropy variables from (2), and the estimates in column (II) by excluding variables related to the installer's rating entropy from (2). Table 3 identifies three key results.

First, the set of variables representing "noise" or dispersion of ratings have a significant impact on an installer's activity level in the marketplace. This is because all entropy variables are found to be statistically significant in the column (III) of Table 3.

Second, an installer's rating dispersion has a positive and statistically significant first-order effect on the installer's activity level because in the column (III), the variable "Rating_Entropy_Self" is found to be significant and its coefficient is positive ($\beta_1 = 1.890, p < 0.001$). On the other hand, we also find that an installer's rating dispersion has a negative and statistically significant second-order effect on its activity level. This is because in the column (III), the variable "Rating_Entropy_Self^2"

is significant and its coefficient is negative ($\beta_2 = -3.473, p < 0.001$). Combining these two effects, an installer's rating dispersion has a concave and non-monotone impact on its activity level in the online marketplace. Specifically, an installer's rating dispersion increases its activity level if and only if the aforementioned dispersion is below a certain threshold; otherwise, any additional dispersion in the installer's ratings lowers its activity level in the marketplace. This finding supports Hypothesis 1A when the installer's rating entropy is smaller than the threshold; otherwise, our finding is in support of Hypothesis 1B.

Third, our estimation shows that the entropy of competitors' ratings impacts an installer's activity level in the similar way as the entropy of the installer's ratings. Specifically, it has a positive and significant first-order effect (as "Rating_Entropy_Others" is significant and its coefficient $\beta_3 = 0.488$ ($p < 0.01$)), and a negative and significant second-order effect (as "Rating_Entropy_Others^2" is significant and its coefficient $\beta_4 = -2.593, p < 0.001$). Combining these two effects, the dispersion in competitors' ratings increases the installer's activity level if and only if the aforementioned dispersion is below a threshold. When the dispersion of competitors' ratings is above that threshold, any additional dispersion in competitors' ratings lowers the installer's marketplace activity. This implies support for Hypothesis 2A if and only if the competitors' rating entropy is below the threshold; otherwise, our findings offer support for Hypothesis 2B.

Figures 3 and 4 illustrate the explained nonlinear effects of the rating entropy on the installer's activity level in the online marketplace. In generating Figures 3 and 4, we use the estimated regression coefficients in the column (III) of Table 3. As is apparent from these figures, on average, the installer's activity level first increases and then decreases with its rating entropy (or the rating entropy of its competitors), yielding an inverted U-shaped relationship between the two.

Apart from these, we also conducted the Shapley decomposition of R-Squared based on Table 3-(III) estimates. Our analysis reveals that the share of installer's rating entropy is more than two times the share of the competitors' rating entropy in explained variance. Thus, the former entropy has a significantly larger marginal contribution to R-square than the latter.

Finally, in all three columns of Table 3, the installer's average rating is significantly and negatively linked with its activity level. Put another way, installers appear to extend fewer proposals as their average ratings increase. One reason for this behavior could be that installers become more selective after they attain a high average rating in the marketplace. Selectiveness can emerge because the installers might think that with a higher average rating, their proposals are more likely to be accepted by customers, and thus making too many offers increases their likelihood of coming across with a negative customer.

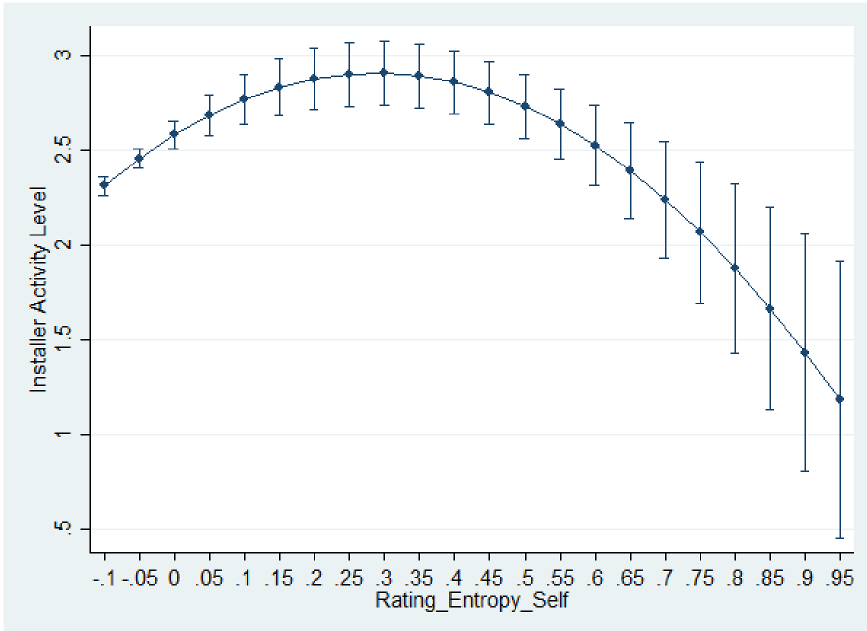

Figure 3 Margins plot for installer's rating entropy versus its activity level with 95% confidence interval

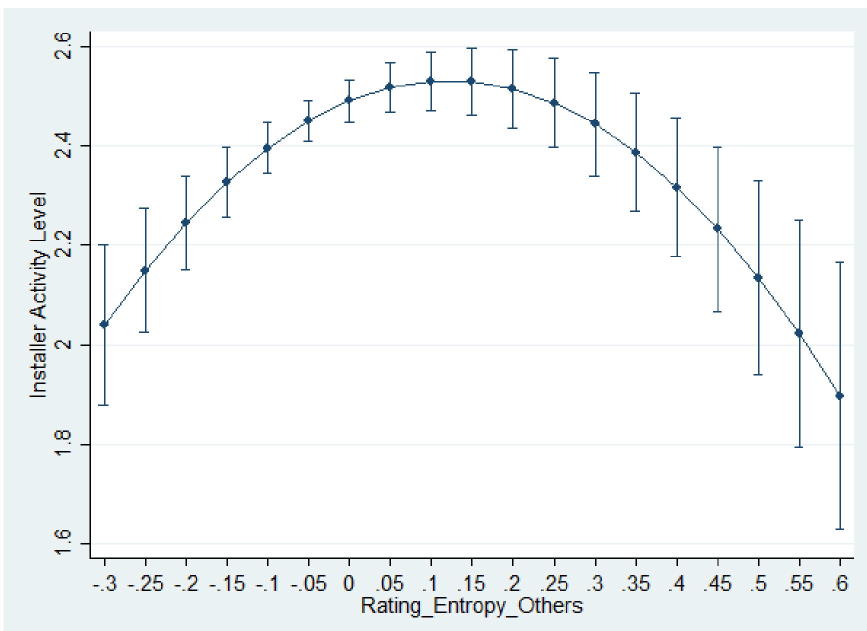

Figure 4 Margins plot for other installer's rating entropy versus the installer's activity level with 95% confidence interval

5. Market-Level Analysis & Results

An important performance metric for the marketplace operator is the number of matches (i.e., agreed proposals) between installers and customers in the marketplace. This section estimates how the market-level rating dispersion impacts *market transaction* that is defined as the logged number of matches in the market. With this, we test Hypotheses 3A and 3B in Section 2.

We will only use numerical ratings in this section. Later, Section 6 will account for text reviews in the market-level analysis. We will provide various additional robustness checks of our findings, and address potential endogeneity concerns in an extended model in Section EC.2.

Recalling that markets and months are indexed by m and t, respectively, we use the following regression equation for the estimation:

$$\begin{aligned}\text{Market\_Transaction}_{m,t+1} = \beta_5 + \beta_6 \text{Rating\_Entropy\_Mkt}_{m,t} + \beta_7 \text{Rating\_Entropy\_Mkt}^2_{m,t} \\ + \text{Controls}_{m,t} + \xi_m + \epsilon_{m,t+1}. \end{aligned} \tag{3}$$

Here, ξ_m is a market-level fixed effect, and it represents the time-invariant market-specific factors that may influence the market transaction.

Performing (3) requires us to convert the installer-level monthly panel data to the market-level monthly panel data based on the markets defined in Section 3.2. Our data include the number of agreed proposals for each installer i in each month t. To create our dependent variable $\text{Market\_Transaction}_{m,t+1}$, we first calculate the total number of proposals accepted by customers in market m and month $t+1$, and then take the natural logarithmic transformation of that sum. Formally, $\text{Market\_Transaction}_{m,t+1} = \ln\left(\sum_{i \in \text{Market } m} \text{Successful\_Proposals}_i + 1\right)$ in (3). We employ this standard transformation because the number of matches is right-skewed and the transformation increases the normality of errors. (As a robustness check, we also perform the analysis without log transformation and the results are consistent.)

A key explanatory variable in (3) is $\text{Rating\_Entropy\_Mkt}_{m,t}$, which is the entropy of all installers' ratings up to (and including) month t in the market m. Because we have two competing hypotheses about the impact of market-level rating entropy (i.e., Hypotheses 3A and 3B in Section 2), we allow for a nonlinear relationship between the market-level rating entropy and the dependent variable. Thus, (3) also contains the quadratic term $\text{Rating\_Entropy\_Market}^2_{m,t}$. The aforementioned two variables are our main explanatory variables, and β_6 and β_7 are the key coefficients of interests. The values of these coefficients together with the significance of the associated explanatory variables will help us determine how the market-level rating entropy impacts market transactions.

In (3), ϵ is the market-level error term, and represents random factors that are unobservable in the data and affect market transactions. We also use various control variables ($\text{Controls}_{m,t}$) in (3). We control for the state of the market. To do that, we created 33 state dummies to represent 33 different states included in the data set. In our dataset, 18% of markets span across more than one state. In light of this, each state dummy represents the fraction of installers that are located in that state within the market m. For example, suppose market 1 has 25% of installers from state X and 75% from Y. Then, we assign 0.25 to the dummy variable State_X and 0.75 to the dummy variable State_Y for market 1. Similar to the installer-level analysis in Section 4, we created the variable Average_Experience$_{m,t}$ that represents the average experience of installers in the market m up to and including month t. In parallel to the installer-level analysis, we use the variable Average_Rating_Mkt$_{m,t}$ to control for the average rating of all installers in the market m until (and including) month t. We also control for the difference between the average unit price of installed 1 KW solar system in the marketplace and off-marketplace, which is represented by the variable Price_Difference_Mkt$_{m,t}$. Finally, we use Market_LogRevenue$_{m,t}$ as a control where it is as defined in Section 4. Summary statistics can be found in Table 4; the correlation coefficients among variables are presented in Table 5. We also checked VIF scores of variables, and verified that they are all within the suggested range (Hair et al. (2014)).

Variables	N	Mean	Standard Deviation	Min	Max
Rating_Entropy_Mkt	642	0	0.236	-0.191	0.809
Market_LogRevenue	642	7.887	8.102	0	22.3
Average_Rating_Mkt	642	4.870	0.245	3	5
Average_Experience	642	1.426	1.110	0	3.332
Price_Difference_Mkt	642	-0.0106	0.146	-0.504	1.312

Table 4 Summary Statistics - Market Level Analysis

Variables	(1)	(2)	(3)	(4)	(5)
(1) Rating_Entropy_Mkt	1				
(2) Average_Rating_Mkt	-0.624	1			
(3) Average_Experience	0.198	-0.107	1		
(4) Price_Difference_Mkt	-0.018	-0.041	-0.012	1	
(5) Market_LogRevenue	0.157	-0.041	0.498	-0.126	1

Table 5 Correlation Matrix - Market Level Analysis

5.1. Results

Table 6 presents our regression estimates for two specifications. Column (I) shows the estimates obtained with the regression (3) in the absence of Rating_Entropy_Mkt$_{m,t}$ and Rating_Entropy_Mkt$^2_{m,t}$ variables, while column (II) includes the estimates obtained by running the regression (3) considering all variables in (3).

Regression estimates reveal the following key findings. First, market-level rating dispersion has a significant and positive first-order effect on the market transaction. This is because the coefficient of "Rating_Entropy_Mkt" is positive ($\beta_6 = 1.060$) and statistically significant ($p < 0.001$) in the column (II) of Table 6. The market-level rating dispersion also has a significant and negative second-order effect on the market transaction as the coefficient of the quadratic term "Rating_Entropy_Mkt^2" is

Variables	(I) Market Transaction	(II) Market Transaction
Rating_Entropy_Mkt		1.060***
		(0.000)
Rating_Entropy_Mkt^2		-1.610***
		(0.000)
Average_Rating_Mkt	-0.273	-0.185
	(0.071)	(0.212)
Average_Experience	0.0197*	0.0138
	(0.013)	(0.073)
Price_Difference_Mkt	0.239	0.287
	(0.282)	(0.194)
Market_LogRevenue	-0.0846	-0.0663
	(0.057)	(0.123)
Constant	3.443**	2.909*
	(0.009)	(0.025)
Market Fixed Effect	Yes	Yes
Weighted State Dummies	Yes	Yes
Observations	642	642
Adjusted R^2	0.739	0.747
AIC	1075.6	1059.6
BIC	1156.0	1148.9

Note: p-value in parentheses; $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$

Table 6 Market Level Analysis

negative ($\beta_7 = -1.160$) and statistically significant ($p < 0.001$) in the column (II). Combining these two effects, regression estimates indicate a concave and non-monotone relationship between the market-level rating dispersion and the market transaction. Specifically, our findings indicate that the market-level rating dispersion increases the market transaction if and only if the mentioned dispersion is smaller than a threshold, for any dispersion beyond that threshold, an increase in the market-level rating dispersion dampens the market transaction (and number of matches). These findings support Hypothesis 3A if and only if the market-level rating dispersion is below a certain threshold; otherwise, our results are in support for Hypothesis 3B.

Figure 5 further illustrates this nonlinear relationship via a margins plot using coefficients generated from estimates in the column (II) of Table 6. As we observe from this figure, on average, the market transaction first increases then decreases with the market-level rating dispersion. Same effect is also valid for the number of matches.

Finally, market-level estimates in Table 6 suggest that after controlling for market conditions, installer experience, price, and state, the average rating is not significantly associated with the market transaction.

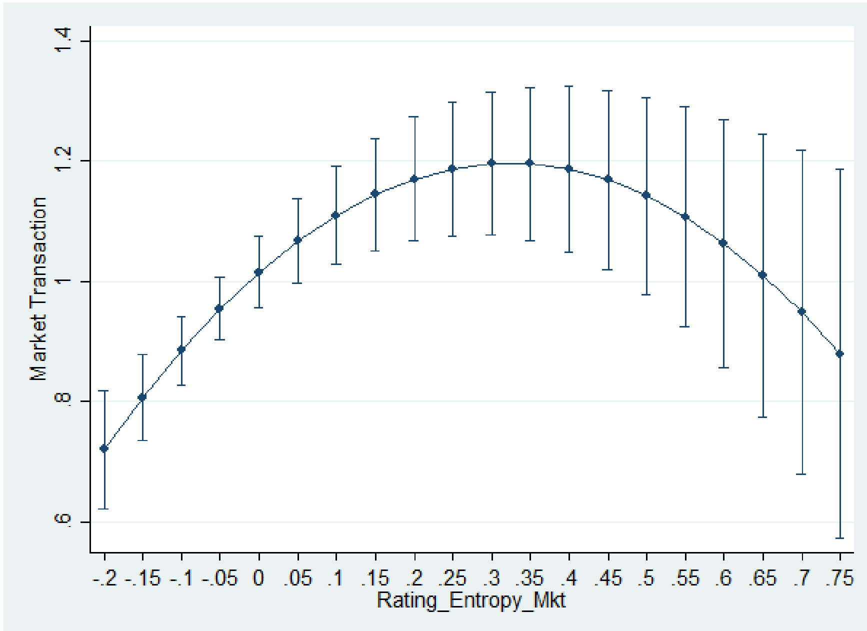

Figure 5 **Margins plot for the market-level rating entropy versus the market transaction with 95% confidence interval**

6. Text Mining

In this section, we incorporate various methods to leverage rich text information in reviews. First, we will use the state-of-the-art advanced natural language processing (NLP) technique called BERT (i.e., *Bidirectional Encoder Representations from Transformers*) to measure the dispersion in text reviews. The BERT method was developed by Google AI in 2018, and incorporated by Google Search Engine in late 2019 (Devlin et al. 2018, Abril 2019). To the best of our knowledge, there is no other paper in operations management literature that considers this deep learning based technique. Second, we will use another state-of-the-art NLP method to generate the sentiment score of each review. As the final step, we will incorporate these text-based metrics in our regression analysis to derive insights related to the text content.

6.1. Measuring Text-Based Review Dispersion via Natural Language Processing Technique BERT

In our data set, we have 3607 pieces of text reviews, and we apply the following five steps to measure the dispersion in text reviews. First, we use BERT to convert each text review to a semantics-sensitive numerical vector. We will provide more details about this method later in this section. As a second step, we normalize each vector to unit length. Third, we measure the cosine similarity between any two review vectors to find the context similarity between the two. The cosine similarity between two normalized vectors V_1 and V_2 equals the inner product of two, i.e., $V_1 \cdot V_2$, and gives the cosine of the angle between the two vectors. The angle represents the similarity in the orientation of two review vectors. For example, if the angle is 0, the vectors are at the same orientation and hence the similarity is the maximum. The second and third steps above are standard in identifying the similarity between two vectors (see, e.g., Hoberg and Phillips (2016)). As the fourth step, we identify the cosine distance between every two review vectors using the fact that the cosine distance between the two normalized vectors V_1 and V_2 equals 1 minus the cosine similarity between the two. This distance reflects how much two reviews differ from each other. As the fifth and final step, we calculate the dispersion of text reviews, i.e., *text-based dispersion*, for a review set of interest by enumerating all pairwise cosine distances of reviews in that set and taking their statistical median (the 50^{th} percentile). For example, for a set of 10 text reviews, we have 45 $(=\binom{10}{2})$ pairwise cosine distances. Finding the text-based dispersion for this set requires computing the median of these 45 distances. If these 10 pieces of texts are dissimilar from each other, they contain richer information and the median of these 45 distances shall be higher; and vice versa. (Section EC.2.4 shows that our findings remain to be valid when the mean (rather than median) of cosine distances is considered in the fifth step of text-based dispersion measurement.)

As a result of this procedure, similar to the rating entropy, we create the following three variables, each measuring the text-based dispersion in a different dimension: (i) $\text{Text\_Dispersion\_Self}_{i,t}$: Demeaned dispersion in installer i's own text reviews up to and including month t. (ii) $\text{Text\_Dispersion\_Others}_{i,t}$: Demeaned dispersion in text reviews of all other installers in the installer i's market up to and including month t. (iii) $\text{Text\_Dispersion\_Mkt}_{m,t}$: Demeaned dispersion in text reviews of all installers in the market m up to and including month t. Note that each of these variables is centered around its mean. We apply this standard procedure because in addition to these terms, we will also consider quadratic terms in our regression.

We now elaborate the BERT model we used to *vectorize* the text reviews. BERT is a natural language processing (NLP) model that transforms texts into numeric vectors while also preserving the meaning of texts. It belongs to the category of NLP methods that performs word embedding. In literature, in different contexts than ours, text data are commonly vectorized based on word counts, ignoring the semantics and word ordering (see, e.g., Hoberg and Phillips (2016) and Loughran and McDonald (2011)). However, our context involves texts that are informal writings and often contain emotions. Simply capturing word frequencies does not provide accurate results if similar emotions can be expressed with synonymous words. Thus, our analysis requires a vectorization that preserves the information and sentiment of the text reviews despite the use of synonyms and/or different styles. The BERT model achieves that. Online Appendix EC.1.1 further explains distinctive features of this method.

6.2. Sentiment Scores of Text Reviews

We use the VADER model to generate sentiment scores for text reviews. VADER is short for *Valence Aware Dictionary and sEntiment Reasoner* and developed by Hutto and Gilbert (2014) as a "parsimonious rule-based model for sentiment analysis of social media text." Since review text shares many structural similarities with the social media text, an important application area of this model is the text analysis of reviews. For each text review, VADER produces a sentiment intensity score from -1 to 1, with 1 representing very positive and -1 representing very negative sentiments.

VADER has key advantages. In contrast to models that use a polarized lexicon where a word is classified as either positive, negative or neutral, VADER is sensitive to both polarity and strength of the sentiment. The method also understands conventional syntactical and grammatical components in the text and reflects them in the sentiment intensity score it generates. Among other features, the method accounts for the exclamation mark, capitalization especially the usage of all-caps, degree adverbs such as "extremely" and "marginally", the contrastive conjunction (e.g.,"but"), conventional emojis, slangs and emoticons in its sentiment intensity score calculation. Example sentiment intensity scores generated with VADER can be found in Online Appendix EC.1.2.

Based on this method, we created three variables that represent average sentiment intensity scores in three dimensions: (i) $\text{Average\_Sentiment\_Self}_{i,t}$: The average sentiment intensity score of installer i's all text reviews up to and including month t. (ii) $\text{Average\_Sentiment\_Others}_{i,t}$: The average sentiment intensity score of competitors' text reviews up to and including month t in the installer i's market. (iii) $\text{Average\_Sentiment\_Mkt}_{m,t}$: The average sentiment intensity score of all text reviews in the market m up to and including month t.

Variables	N	Mean	Standard Deviation	Min	Max
Text_Dispersion_Self	4,562	0	0.0711	-0.064	0.318
Text_Dispersion_Others	4,562	0	0.0219	-0.061	0.241
Average_Sentiment_Self	4,562	0.411	0.358	-0.827	0.958
Average_Sentiment_Others	4,562	0.561	0.283	-0.599	0.958
Text_Dispersion_Mkt	642	0	0.036	-0.062	0.235
Average_Sentiment_Mkt	642	0.463	0.334	0.0593	0.866

Table 7 **Summary statistics of text-based variables for installer and market-level analysis**

6.3. Empirical Analysis Using Variables Derived From Text Mining

We now discuss the analysis we conducted with the text-based dispersion and average sentiment scores we constructed in Sections 6.1 and 6.2. With these additional variables, we aim to examine the following questions: (i) Is the text content significant in explaining installers' activity levels and market transactions? (ii) How do the text-based dispersion and the average sentiment intensity score influence an installer's activity level and market transactions? To study these questions, we consider the following regression models:

$$
\begin{aligned}
&\text{Installer\_Activity}_{i,m,t+1} \\
&= \theta_0 + \theta_1 \text{Rating\_Entropy\_Self}_{i,t} + \theta_2 \text{Rating\_Entropy\_Self}^2_{i,t} + \theta_3 \text{Rating\_Entropy\_Others}_{i,t} \\
&+ \theta_4 \text{Rating\_Entropy\_Others}^2_{i,t} + \theta_5 \text{Text\_Dispersion\_Self}_{i,t} + \theta_6 \text{Text\_Dispersion\_Self}^2_{i,t} \\
&+ \theta_7 \text{Text\_Dispersion\_Others}_{i,t} + \theta_8 \text{Text\_Dispersion\_Others}^2_{i,t} + \text{Controls}_{i,m,t} + \alpha_i + \epsilon_{i,t+1}. \qquad (4)
\end{aligned}
$$

$$
\begin{aligned}
&\text{Market\_Transaction}_{m,t+1} \\
&= \eta_0 + \eta_1 \text{Rating\_Entropy\_Mkt}_{m,t} + \eta_2 \text{Rating\_Entropy\_Mkt}^2_{m,t} \\
&+ \eta_3 \text{Text\_Dispersion\_Mkt}_{m,t} + \eta_4 \text{Text\_Dispersion\_Mkt}^2_{m,t} + \text{Controls}_{m,t} + \xi_m + \epsilon_{m,t+1}. \qquad (5)
\end{aligned}
$$

In (4) and (5), all variables except text-based variables and control variables on average rating are the same as the ones in (2) and (3), respectively. An installer's text-based dispersion correlates with its rating entropy on a lower level (correlation coefficient is 0.304; see Tables EC.1 and EC.2 for the entire correlation matrix). Thus, (4) and (5) consider the text-based dispersion and rating entropy variables in the same regression. On the other hand, by Table EC.1, an installer's average sentiment intensity score significantly correlates with its average rating (with a correlation coefficient of 0.832). As a result, in (4) and (5), we will use average sentiment intensity scores and average rating variables as substitute controls. Table EC.4 shows the estimation results obtained by using different set of explanatory variables in (4). In each column of Table EC.4, blank represents the absence of the corresponding explanatory variable in the regression.

Regression estimates in Table EC.4 reveal key findings. First, an installer's or its competitors' rating entropy continues to have an inverted U-shaped impact on the installer's activity level even when the corresponding text-based variables are also considered. Second, an installer's text-based dispersion has a significant and positive first-order effect while having a significant and negative second-order effect on the installer's activity level (i.e., $\theta_5 > 0$ and $\theta_6 < 0$ are both significant). Combining the two, an installer's text-based dispersion has an inverted U-shaped impact on its activity level. In contrast, competitors' text-based dispersion does not have a significant impact on the installer's activity level. One explanation for this result can be that installers invest more time in understanding the text content of their own reviews than others' reviews. In fact, from our discussions with practitioners we learned that installers read their own text reviews very carefully. However, installers can conceivably prioritize ratings over text reviews in evaluating the competitors' review dispersion because (i) on average, the total number of competitors' reviews is much higher than the number of installer's reviews, and (ii) the average length per text review is long (558 characters), implying that evaluating each text review is a much harder task than considering a numerical rating.

Table EC.5 displays the estimates under different specifications of the market-level model. There are three key findings. First, the market-level rating entropy continues to be significant and have an inverted U-shaped impact on market transactions (and number of matches). Thus, our original finding in Section 5 is robust. Second, although the text-based dispersion has an inverted U-shaped relationship with market transactions, its impact is statistically insignificant. This contrast between rating- and text-based dispersions, albeit both measuring dispersion, can be explained as follows. Recall that the response variable in (5) is the logged number of matches between customers and installers. Given the length and volume of text reviews, customers could pay less attention to the text-based review dispersion than to the rating-based dispersion as the former requires deeper evaluation of reviews. Therefore, a text-based measure can have less significance than a rating-based measure in explaining the impact of review dispersion on the market transaction. Finally, our third finding is that the average market-level sentiment score has a significant and negative impact on the market transaction. This sign is consistent with its counterparts in the rating-based analysis.

7. Conclusions

Our paper contributes to the literature by empirically investigating if and how the dispersion in customer reviews impacts a firm's activity level and the number of matches in an online marketplace where firms are active. To the best of our knowledge, there is no prior work that examines this topic. Our findings offer key insights to a marketplace operator.

We find that there is a significant and inverted U-shaped relationship between a firm's review dispersion and its activity level in the online marketplace. Thus, an increase in a firm's review dispersion can increase or decrease its activity level in the marketplace, depending on the level of dispersion. If the mentioned dispersion is below (respectively, above) a certain threshold, an increase in that dispersion increases (respectively, decreases) the firm's activity level in the marketplace. Furthermore, we find that a firm's activity level has a significant and inverted U-shaped relationship with competitors' rating dispersion in the online marketplace. Thus, similar to the previous finding, an increase in this type of dispersion can encourage or deter the firm from making a proposal, depending on the dispersion level.

We identify a significant and inverted U-shaped relationship between the market transactions and the review dispersion at a local market level. This finding has a key implication for an online marketplace operator: Having all sellers with 5 stars might not be favorable to the marketplace operator. Review dispersion up to a particular level can help an online marketplace operator in terms of number of matches.

Regarding the methodology, to analyze text reviews, we incorporated two text-mining methods: VADER that assigns a one-dimensional sentiment intensity score to each text review and BERT that converts each piece of text review to a numerical vector via deep learning. We used the latter to measure the content dissimilarity among text reviews with precision. To our knowledge, our paper is the first that uses the deep-learning based advanced text-mining method BERT in the operations management literature. Apart from this, our paper provides a showcase for the clustering method OPTICS that has many advantages over common clustering techniques. This advanced clustering method has not been used in the OM literature yet. Overall, these methods have the potential to facilitate research in various contexts in the operations management literature.

References

Abril D (2019) Google Says Its Latest Tech Tweak Provides Better Search Results. Here's How. *Fortune* `fortune.com/2019/10/25/google-tweaks-search-results-ai/`, Accessed on 6/2/20.

Ali F (2020) What are the top online marketplaces? `www.digitalcommerce360.com/article/infographic-top-online-marketplaces/`, Accessed on 1/11/20.

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies* 58(2):277–297.

Askalidis G, Malthouse EC (2016) The value of online customer reviews. *Proceedings of the 10th ACM Conference on Recommender Systems*, 155–158.

Baker S (2020) Investing in online marketplaces. *Scale Venture Partners* `https://www.scalevp.com/blog/investing-in-online-marketplaces`, Accessed on 6/15/20.

Baumol WJ, Ide EA (1956) Variety in retailing. *Management Science* 3(1):93–101.

Bimpikis K, Elmaghraby WJ, Moon K, Zhang W (2019) Managing market thickness in online B2B markets Forthcoming in Management Science.

Blanco CC, Caro F, Corbett CJ (2020) Do carbon abatement opportunities become less profitable over time? A global firm-level perspective using CDP data. *Energy Policy* 138:111252.

Blass V, Corbett CJ, Delmas MA, Muthulingam S (2014) Top management and the adoption of energy efficiency practices: Evidence from small and medium-sized manufacturing firms in the US. *Energy* 65:560–571.

Boris (2018) How we determine valuations for marketplaces. `https://versionone.vc/how-we-determine-valuations-for-marketplaces/`, Accessed on 6/15/20.

Chakraborty U, Bhat S (2018a) Credibility of online reviews and its impact on brand image. *Management Research Review* 41(1):148–164.

Chakraborty U, Bhat S (2018b) The effects of credible online reviews on brand equity dimensions and its consequence on consumer behavior. *Journal of Promotion Management* 24(1):57–82.

Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research* 43(3):345–354.

Chintagunta PK, Gopinath S, Venkataraman S (2010) The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science* 29(5):944–957.

Clemons EK, Gao GG, Hitt LM (2006) When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of Management Information Systems* 23(2):149–171.

Davidson M (2019) What Neighborhood Is This? Estimating Material Based Neighborhood Boundaries Using Single-Family Homes Working paper, University of Michigan at Ann Arbor.

Delfgaauw J, Dur R, Sol J, Verbeke W (2013) Tournament Incentives in the Field: Gender Differences in the Workplace. *Journal of Labor Economics* 31(2):305–326.

Demirag I, Khadaroo I, Stapleton P, Stevenson C (2011) Risks and the financing of PPP: Perspectives from the financiers. *The British Accounting Review* 43(4):294–310.

Devlin J, Chang MW, Lee K, Toutanova K (2018) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding `https://arxiv.org/abs/1810.04805`, Accessed on 6/15/20.

Dhanorkar S, Donohue K, Linderman K (2015) Repurposing Materials and Waste through Online Exchanges: Overcoming the Last Hurdle. *Production and Operations Management* 24(9):1473–1493.

Duan W, Gu B, Whinston AB (2008) Do online reviews matter?An empirical investigation of panel data. *Decision support systems* 45(4):1007–1016.

EnergySage (2019) New Report Finds Solar Tariffs Not Helping American-Made Solar Panels Compete, and Costs of Solar Continuing to Fall. `https://www.energysage.com/press/energysage-marketplace-intel-report-8`, Accessed on 1/20/20.

Galston E (2017) Anatomy of a managed marketplace. `https://techcrunch.com/2017/05/25/anatomy-of-a-managed-marketplace/`, Accessed on 6/15/20.

Gensler S, Vlckner F, Egger M, Fischbach K, Schoder D (2015) Listen to Your Customers: Insights into Brand Image Using Online Consumer-Generated Product Reviews. *International Journal of Electronic Commerce* 20(1):112–141.

Hair JF, Black JWC, Babin BJ, Anderson RE (2014) Multivariate Data Analysis Seventh Edition .

Hirshleifer D, Thakor AV (1992) Managerial conservatism, project choice, and debt. *The Review of Financial Studies* 5(3):437–470.

Hoberg G, Phillips G (2016) Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124(5):1423–1465.

Huang H, Narayanan S, Swaminathan JM (2020) Carrot or Stick? Supplier Diversity and its Impact on Carbon Emission Reduction Strategies Working Paper, UNC Kenan-Flagler Business School.

Hutto CJ, Gilbert E (2014) Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Eighth international AAAI conference on weblogs and social media.*

IEA (2018) Snapshot of global photovoltaic markets 2018 `http://www.iea-pvps.org/fileadmin/dam/public/report/statistics/IEA-PVPS_-_A_Snapshot_of_Global_PV_-_1992-2017.pdf`, Accessed on 6/15/20.

Iyengar SS, Lepper MR (2000) When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and social psychology* 79(6):995.

Kanagala HK, Krishnaiah VJR (2016) A comparative study of k-means, dbscan and optics. *2016 International Conference on Computer Communication and Informatics (ICCCI)*, 1–6 (IEEE).

Kesavan S, Staats BR, Gilland W (2014) Volume flexibility in services: The costs and benefits of flexible labor resources. *Management Science* 60(8):1884–1906.

Kundu A, Ramdas K (2019) Timely After-Sales Service and Technology Adoption: Evidence from the Off-Grid Solar Market in Uganda Working paper, London Business School.

Lee HL, Tang CS (2018) Socially and Environmentally Responsible Value Chain Innovations: New Operations Management Research Opportunities. *Management Science* 64(3):983–996.

Li J, Netessine S (2020) Higher market thickness reduces matching rate in online platforms: Evidence from a quasiexperiment. *Management Science* 66(1):271–289.

Lind JT, Mehlum H (2010) With or without U? The appropriate test for a U-shaped relationship. *Oxford bulletin of economics and statistics* 72(1):109–118.

Loughran T, McDonald B (2011) When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance* 66(1):35–65.

Luca M (2016) Reviews, reputation, and revenue: The case of Yelp. com Working Paper, Harvard Business School.

Luo X, Raithel S, Wiles MA (2013) The impact of brand rating dispersion on firm value. *Journal of Marketing Research* 50(3):399–415.

Moreno A, Terwiesch C (2014) Doing business with strangers: Reputation in online service marketplaces. *Information Systems Research* 25(4):865–886.

Nickell S (1981) Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society* 1417–1426.

Niewoehner III RJ, Staats BR (2020) Focusing Provider Attention: An Empirical Examination of Incentives and Feedback in Flu Vaccinations Working paper, UNC Kenan-Flagler Business School.

Park DH, Lee J, Han I (2007) The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement. *International Journal of Electronic Commerce* 11(4):125–148.

Patel N (2020) How to Get More Online Reviews (and Deal With Bad Ones) `https://neilpatel.com/blog/how-to-get-more-online-reviews-and-deal-with-bad-ones/`, Accessed on 6/6/20.

Pyper J (2018) California's rooftop solar mandate wins final approval. `https://www.greentechmedia.com/articles/read/california-solar-roof-mandate-wins-final-approval`, Accessed on 6/15/20.

Scheibehenne B, Greifeneder R, Todd PM (2010) Can there ever be too many options? A meta-analytic review of choice overload. *Journal of consumer research* 37(3):409–425.

SEIA (2019) Solar industry research data. URL `https://www.seia.org/solar-industry-research-data`.

Singh SP, Scheller-Wolf A (2018) That's Not Fair: Tariff Structures for Electricity Markets with Rooftop Solar Working Paper, UCL School of Management.

Smaldino PE (2013) Measures of individual uncertainty for ecological models: Variance and entropy. *Ecological modelling* 254:50–53.

Song H, Tucker AL, Murrell KL, Vinson DR (2017) Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. *Management Science* 64(6):2628–2649.

Sunar N, Birge JR (2019) Strategic Commitment to a Production Schedule with Uncertain Supply and Demand: Renewable Energy in Day-Ahead Electricity Markets. *Management Science* 65(2):714–734.

Sunar N, Swaminathan JM (2018) Net-Metered Distributed Renewable Energy: A Peril for Utilities? Working Paper,UNC Kenan-Flagler Business School.

Tan TF, Netessine S (2014) When does the devil make work? An empirical study of the impact of workload on worker productivity. *Management Science* 60(6):1574–1593.

Tan TF, Netessine S (2019) When you work with a superman, will you also fly? An empirical study of the impact of coworkers on performance. *Management Science* 65(8):3495–3517.

Teimouri M, Indahl UG, Tveite H (2016) A method to detect inactive periods in animal movement using density-based clustering. *Applied Geography* 73:102–112.

Spielger Research Center (2017) How Online Reviews Influence Sales `https://spiegel.medill.northwestern.edu/_pdf/Spiegel_Online\%20Review_eBook_Jun2017_FINAL.pdf`, Accessed on 6/2/20.

US Census Bureau (2019) Gazetteer files. `https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.2018.html`, Accessed on 6/15/20.

US EIA (2020a) Annual Energy Outlook 2020 with projections to 2050 `https://www.eia.gov/outlooks/aeo/pdf/AEO2020\%20Full\%20Report.pdf`, Accessed on 6/11/20.

US EIA (2020b) Energy Power Monthly - Table 1.1.A. Net Generation from Renewable Sources: Total (All Sectors), 2010-January 2020 `https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=epmt_1_01_a`, Accessed on 6/11/20.

US EIA (2020c) Form EIA-861m (formerly EIA-826) detailed data - Estimated Small Scale Solar PV Capacity `www.eia.gov/electricity/data/eia861m/`, Accessed on 1/11/20.

Uppari B, Netessine S, Popescu I, Barron M, Clarke R, Visser M (2019) Inconvenience, Liquidity Constraints, and Adoption of Off-Grid Lighting Solutions Working Paper, Singapore Management University.

Weaver J (2019) Residential solar power growing like a "weed", straining labor. https://pv-magazine-usa.com/2019/10/28/residential-solar-power-growing-like-a-weed-straining-labor/, Accessed on 6/15/20.

Zhang X (2006) *Tapping into the pulse of the market: Essays on marketing implications of information flows.* Ph.D. thesis, Massachusetts Institute of Technology.

Zhu F, Zhang XM (2010) Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of Marketing* 74(2):133–148.

Electronic Companion for "Do Noisy Customer Reviews Discourage Platform Sellers? Empirical and Textual Analysis of an Online Solar Marketplace Using Deep Learning"

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Average_Sentiment_Self	1							
(2) Average_Rating_Self	0.832	1						
(3) Average_Sentiment_Other	-0.036	-0.074	1					
(4) Average_Rating_Other	-0.073	-0.023	0.434	1				
(5) Text_Dispersion_Other	-0.064	0.005	-0.032	-0.076	1			
(6) Rating_Entropy_Other	0.122	0.07	-0.089	-0.527	0.081	1		
(7) Text_Dispersion_Self	0.1	0.136	-0.089	-0.003	-0.091	0.009	1	
(8) Rating_Entropy_Self	-0.081	-0.086	0.064	0.023	0.026	-0.089	0.304	1

Table EC.1 **Correlation of Ratings and Text-based Measures (Installer-level)**

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Average_Sentiment_Self	1.000							
(2) Average_Sentiment_Others	-0.024	1.000						
(3) Text-based_Entropy_Others	-0.063	-0.055	1.000					
(4) Text-based_Entropy_Self	0.106	-0.022	-0.092	1.000				
(5) Review_Count	0.166	0.070	-0.027	0.241	1.000			
(6) Experience	0.082	0.205	-0.064	0.082	0.124	1.000		
(7) Price_Difference	-0.039	0.005	-0.009	-0.048	-0.027	-0.033	1.000	
(8) Market_LogRevenue	0.034	0.246	-0.127	0.003	-0.052	0.553	-0.062	1.000

Table EC.2 **Correlation Matrix - Individual Level Text-based Analysis**

EC.1. Additional Information and Tables for Section 6

EC.1.1. More details about the BERT method

The BERT model has two distinct advantages over existing methods. First, it understands the semantics. For example, consider the 3 sentences:

Sentence 1: `they did a good job.` Sentence 2: `they did an awful job.` Sentence 3: `they did a great job.`

Considering the meaning of the sentences, we expect the distance between sentences 1 and 3, $D(1,3)$, to be smaller than the distance between 2 and 3 or 1 and 2, i.e., $D(2,3)$ or $D(1,2)$. The BERT model vectorization enables just that; it projects "good" and "great" to vectors that are closer to each other. In this example, with BERT, we have $D(1,3) = 0.03 < D(1,2) = 0.09 < D(2,3) = 0.1$. This level of distinction is not feasible without word embedding (e.g., by simply using a word counter vectorizer).

Second, the BERT model takes word ordering into account. For example, the two sentences "The food was good, not bad at all" and "The food was bad, not good at all" have the opposite meanings. Common vectorization methods (e.g., "bag-of-words" approach) are not able to capture this difference as words and number of counts are the same in both sentences. But, the BERT model can easily differentiate between these two sentences.

EC.1.2. Examples under the VADER method

The following review, which was rated as 5-star, received a sentiment score of 0.8622 under the VADER method:

"*Mike at (...) was friendly, courteous, professional and very helpful. At first I did not know what kind of system I wanted, because my roof was too small and I had some trees in the way. Mike had never installed a tracking system but he did recommend it. It seemed like we would get the best "bang for the buck" with this system, so I went with it. Mike had all subcontractor there on time as well as all the equipment. It was up and running in less than a week. I love it.*"

As another example, the following review, which was rated as 1-star, received a sentiment score of -0.7184 under the VADER method:

"*Do not hire (...) to install a solar system. Do not hire (...) to do anything. Evan and all his various companies and names ARE NOT LICENCED OR INSURED. I was scammed by Mr. Evan (...) in December of 2013. He installed the system wrong and incomplete even though all the parts and materials were provided for him. Please take the time to do your research and check references and validate licenses and insurance information. It will save u more money than to trust a cheap con artist. All the info at (...) is fraudulent lies. Evan (...) is also known as (...).*"

EC.1.3. Tables for Section 6

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Text_Dispersion_Mkt	1						
(2) Rating_Entropy_Mkt	0.068	1					
(3) Average_Sentiment_Mkt	-0.43	0.118	1				
(4) Average_Rating_Mkt	-0.097	-0.624	0.173	1			
(5) Average_Experience	-0.113	0.198	0.053	-0.107	1		
(6) Price_Difference_Mkt	0.127	-0.018	0.017	-0.041	-0.012	1	
(7) Market_LogRevenue	-0.087	0.157	0.23	-0.041	0.498	-0.126	1

Table EC.3 Correlation Matrix - Market Level Text-based Analysis

Variables	(I) Installer's Activity Level	(II) Installer's Activity Level	(III) Installer's Activity Level	(IV) Installer's Activity Level
Text-based_Entropy_Self	5.326***	5.389***	5.014***	4.890***
	(0.000)	(0.000)	(0.000)	(0.000)
Text-based_Entropy_Self^2	-23.71***	-23.54***	-20.29***	-20.40***
	(0.000)	(0.000)	(0.000)	(0.000)
Text-based_Entropy_Others	-1.288	-0.989	-1.640	-1.955
	(0.389)	(0.509)	(0.275)	(0.193)
Text-based_Entropy_Others^2	-9.593	-13.13	1.348	5.132
	(0.623)	(0.517)	(0.944)	(0.782)
Average_Rating_Self	-0.942***			-0.998***
	(0.000)			(0.000)
Average_Rating_Others	0.000219			-0.00483
	(0.991)			(0.813)
Average_Sentiment_Self		-0.431	-0.404	
		(0.090)	(0.129)	
Average_Sentiment_Others		0.111	0.0806	
		(0.378)	(0.524)	
Rating_Entropy_Self			2.044***	2.161***
			(0.000)	(0.000)
Rating_Entropy_Self^2			-4.246***	-4.344***
			(0.000)	(0.000)
Rating_Entropy_Others			0.399*	0.384*
			(0.033)	(0.045)
Rating_Entropy_Others^2			-2.382***	-2.376***
			(0.000)	(0.000)
Review_Count	0.0489***	0.0492***	0.0420***	0.0413***
	(0.000)	(0.000)	(0.000)	(0.000)
Experience	0.178***	0.175***	0.176***	0.177***
	(0.001)	(0.001)	(0.001)	(0.001)
Price_Difference	0.0952	0.101	0.125	0.118
	(0.278)	(0.247)	(0.158)	(0.181)
Market_LogRevenue	-0.0164***	-0.0165***	-0.0159***	-0.0158***
	(0.000)	(0.000)	(0.001)	(0.001)
Constant	2.374***	2.075***	2.379***	2.649***
	(0.000)	(0.000)	(0.000)	(0.000)
Fixed Effect	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes
Observations	4562	4562	4562	4562
Adjusted-R^2	0.633	0.633	0.638	0.638
AIC	13202.7	13200.9	13147.9	13147.6
BIC	13292.7	13290.8	13263.5	13263.2

Note: p-value in parentheses; $^{\star}p < 0.05$;$^{\star\star}p < 0.01$;$^{\star\star\star}p < 0.001$

Table EC.4 Installer Level Analysis with Variables Derived from Text Analysis

Variables	(I) Market Transaction	(II) Market Transaction	(III) Market Transaction
Rating_Entropy_Mkt			1.605***
			(0.000)
Rating_Entropy_Mkt^2			-1.622***
			(0.000)
Average_Rating_Mkt	-0.259		
	(0.092)		
Average_Sentiment_Mkt		-1.316**	-1.180*
		(0.006)	(0.012)
Average_Experience	0.0193*	0.0148	0.00983
	(0.016)	(0.058)	(0.196)
Price_Difference_Mkt	0.228	0.228	0.292
	(0.314)	(0.301)	(0.182)
Market_LogRevenue	-0.0820	-0.0652	-0.0517
	(0.066)	(0.125)	(0.214)
Text_Dispersion_Mkt	3.014	1.318	-1.856
	(0.443)	(0.736)	(0.632)
Text_Dispersion_Mkt^2	-7.258	-3.138	3.539
	(0.420)	(0.730)	(0.691)
Constant	3.124*	2.843*	2.688*
	(0.024)	(0.017)	(0.021)
Fixed Effect	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes
Observations	642	642	642
Adjusted R^2	0.739	0.743	0.750
AIC	8853.4	8849.3	8843.7
BIC	8925.9	8933.9	8928.3

Note: p-value in parentheses; $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$

Table EC.5 Market Level Analysis with Variables Derived from Text Analysis

EC.2. Robustness Checks

EC.2.1. Dynamic Panel Model

The regression model in Section 4 considers fixed effect for each installer, and that accounts for time-invariant installer-specific factors that may impact the dependent variable, i.e., installer's activity level. This section extends our regression model (2) to a dynamic panel model by including lagged dependent variables. The inclusion of these variables aims to consider any other unobserved heterogeneity that may influence the dependent variable but not captured in (2). In light of this, for the installer-level analysis, our regression equation is extended to the following:

$$
\begin{aligned}
&\text{Installer\_Activity}_{i,m,t+1} \\
&= \gamma_0 + \gamma_1 \text{Installer\_Activity}_{i,t} + \gamma_2 \text{Installer\_Activity}_{i,t-1} + \gamma_3 \text{Rating\_Entropy\_Self}_{i,t} \\
&+ \gamma_4 \text{Rating\_Entropy\_Self}^2_{i,t} + \gamma_5 \text{Rating\_Entropy\_Others}_{i,t} + \gamma_6 \text{Rating\_Entropy\_Others}^2_{i,t}
\end{aligned}
$$

$$+\alpha_i + \text{Controls}_{i,m,t} + \epsilon_{i,t+1}. \tag{EC.1}$$

However, the inclusion of the lagged dependent variables in the presence of fixed effects may cause endogeneity bias, as such an addition may lead a correlation between the regressors and the error (Nickell 1981). To overcome this, we use Arellano and Bond (1991)'s method that addresses endogeneity bias in dynamic panel data. Arellano and Bond (1991) estimator is a general method of moments estimator that is based on dynamic panel data with first differences. It uses lagged variables as instruments to address the endogeneity bias. Arellano and Bond (1991) estimation requires serially uncorrelated first-differenced errors. We provide support for this property in Table EC.6. We also modify the market-level model (3) to include lagged dependent variables:

$$\begin{aligned}
&\text{Market_Transaction}_{m,t+1} \\
&= \eta_0 + \eta_1 \text{Market_Transaction}_{m,t} + \eta_2 \text{Market_Transaction}_{m,t-1} + \eta_3 \text{Market_Transaction}_{m,t-2} \\
&+ \eta_4 \text{Rating_Entropy_Mkt}_{m,t} + \eta_5 \text{Rating_Entropy_Mkt}^2_{m,t} + \xi_m + \text{Controls}_{m,t} + \epsilon_{m,t+1}.
\end{aligned} \tag{EC.2}$$

Here, ξ_m represents the time-invariant market-specific factors that may influence market transaction. Table EC.6 also provides support for serially-uncorrelated first-differenced errors on the market-level. Thus, to overcome any potential endogeneity bias in this regression, we apply similar steps as the ones explained for the analysis of the installer-level activity. In applying Arellano and Bond (1991) estimator, we included $1-3$ lags of variables. We also note that (EC.1) includes two lagged dependent variables while (EC.2) includes three lagged dependent variables simply to be consistent with the requirements of Arellano and Bond (1991).

Tables EC.7 and EC.8 include our installer-level and market-level estimates. An installer's rating entropy continues to have an inverted U-shaped impact on the installer's activity level. Furthermore, there is also an inverted U-shaped relationship between the market-level rating entropy and market transaction (or number of matches). Thus, our key findings are robust in this extension.

EC.2.2. Additional Support for Inverted U-Shaped Relationship

To further validate the inverted U-shaped relationship between an explanatory variable and the response variable, one must check whether the stationary point of the explanatory variable lies within its range in our sample. This check is important to distinguish the inverted U-shaped relationship from a concave monotone relationship. The ranges of "Rating_Entropy_Self," "Rating_Entropy_Others" and "Rating_Entropy_Mkt" are provided in Tables 1 and 4. Based on our estimates in Sections 4 and 5, we calculate the stationary points for these variables as $S_{\text{self}} \doteq -\beta_1/(2\beta_2) = 0.272$, $S_{\text{others}} \doteq -\beta_3/(2\beta_4) = 0.094$ and $S_{\text{mkt}} \doteq -\beta_6/(2\beta_7) = 0.329$. Comparing ranges and the stationary points, we conclude that in our data, the stationary point of each rating entropy variable lies within its observed data range. In fact, stationary point for each of these variables is also evident in Figures 3, 4 and 5. This provides further validation for the inverted U-shaped relationship we find in Sections 4 and 5.

We used a common criteria to identify inverted U-shaped relationships. Some researcher argue that for the inverted U-shaped relationship to be meaningful for an explanatory variable, the stationary point for that variable should not be too close to the end points of the data range or too far from the sample mean (e.g., 3 standard deviation far) (Lind and Mehlum 2010). This concern does not apply to our analysis as the stationary point for each rating entropy is close to its sample mean. Specifically, S_{self}, S_{others} and S_{mkt} are respectively 1.25, 0.51 and 1.39 standard deviation away from their sample means.

EC.2.3. Alternative Test for Inverted U-Shaped Relationship: Spline Regression

Up to this section, we identified an inverted U-shaped relationship between entropy measures and dependent variables by applying a standard technique. That is, by running a polynomial regression, showing the significance of linear and quadratic terms of entropy measure, and identifying the positive sign for the linear term and the negative sign for the quadratic term. See, e.g., Tan and Netessine (2014) and Kesavan et al. (2014) that apply this technique. For robustness check, we also perform spline regressions on both individual and market level analysis to identify the non-monotone effects via a different approach.

Spline regressions use breakpoints (*knots*) to capture the changes in coefficients for different intervals of explanatory variables. We perform spline regressions with 1 knot and with 2 knots for entropy variables. Knots divide the range of the explanatory variable of interest into sub-ranges. For each of these sub-ranges, we create a spline variable of the considered explanatory variable, and allow for a different linear relationship between the response variable and the spline variable. For example, for 1 knot, we create two spline variables Rating_Entropy_Others_1 and Rating_Entropy_Others_2, and consider the linear term of either one only in one of the two ranges of the variable. For the installer-level analysis, we plug in the linear term of Rating_Entropy_Others_1 (respectively, Rating_Entropy_Others_2) in place of the combination of linear and quadratic terms of Rating_Entropy_Others in (2) when Rating_Entropy_Others is smaller (respectively, larger) than the breakpoint. Likewise, in separate regressions, we repeat this procedure for Rating_Entropy_Self in the installer-level analysis (based on (2)) and for Rating_Entropy_Mkt in the market-level analysis (based on (3)). We also further extend this alternative testing to consider spline regressions with 2 knots, which require us to create three spline variables for each rating entropy measure.

The results are presented in Tables EC.9 and EC.10. In Table EC.9, we report the spline regression estimates with one knot on Rating_Entropy_Others in column (I) and with one knot on

Rating_Entropy_Self in column (II). We find that the coefficient of the first spline is positive and significant while the second one (which is valid above the breakpoint) is negative and significant ($p < 0.001$), supporting the inverted U-shaped relationships between either rating entropy measure and the installer's activity level. The conclusions are similar when we consider the case with 2 knots as shown in column (III) and (IV) of Table EC.9. Next, we consider the market-level analysis with results presented in Table EC.10. For the case with 1 breakpoint for Rating_Entropy_Mkt, first positive and then negative and significant ($p < 0.001$) coefficients associated with the two splines in column (I) further validate the inverted U-shaped relationship we established on the market-level. We also find that when we move to the case with two breakpoints, non-monotone relationship is still preserved in the market level.

EC.2.4. Alternative Approach to Measure Text-based Dispersion

In Section 6, we measured the text-based dispersion by taking the median of cosine distances. Alternatively, one can consider the mean of cosine distances to create the text-based dispersion variables. Tables EC.11 and EC.12 report the estimation results when the mean (rather than median) of cosine distances is considered. By Table EC.11, the installer-level text-based dispersion result in Section 6 continues to hold. In addition, Table EC.12 suggests that when the mean (rather than median) of cosine distances is considered, the market-level text-based review dispersion has a significant and inverted U-shaped impact on the market transaction.

EC.3. Tables For Section EC.2

Installer Level Dynamic Panel		
Order	z	$p-value$
H_0: No correlation between $\Delta_{i,t}$ and $\Delta_{i,t-1}$	-9.8283	0.00
H_0: No correlation between $\Delta_{i,t}$ and $\Delta_{i,t-2}$	-0.66053	0.5089
Sargan Test for overidentifying restriction	$\chi^2(1525) = 238.2181$	1.000
Market Level Dynamic Panel		
Order	z	$p-value$
H_0: No correlation between $\Delta_{i,t}$ and $\Delta_{i,t-1}$	-2.7882	0.01
H_0: No correlation between $\Delta_{i,t}$ and $\Delta_{i,t-2}$	0.04295	0.9657
Sargan Test for overidentifying restriction	$\chi^2(544) = 22.77427$	1.000

Table EC.6 Dynamic Panel Specification Checks

Variables	(I) Installer's Activity Level	(II) Installer's Activity Level	(III) Installer's Activity Level	(IV) Installer's Activity Level
Installer_Activity$_t$	0.510***	0.507***	0.509***	0.502***
	(0.000)	(0.000)	(0.000)	(0.000)
Installer_Activity$_{t-1}$	0.0393	0.0374	0.0363	0.0369
	(0.130)	(0.155)	(0.176)	(0.158)
Rating_Entropy_Self			1.322**	1.351**
			(0.006)	(0.004)
Rating_Entropy_Self^2			-1.578*	-1.740*
			(0.043)	(0.026)
Rating_Entropy_Others		0.733		0.695
		(0.059)		(0.069)
Rating_Entropy_Others^2		-1.300*		-1.253*
		(0.041)		(0.031)
Average_Rating_Self	-0.0618***	-0.0713***	-0.0662***	-0.0680***
	(0.000)	(0.000)	(0.000)	(0.000)
Average_Rating_Others	-0.180	-0.191	-0.137	-0.162
	(0.128)	(0.138)	(0.232)	(0.183)
Review_Count	0.0165*	0.0175**	0.0134*	0.0105
	(0.023)	(0.007)	(0.024)	(0.077)
Experience	-0.0554	-0.0445	-0.0564	-0.0528
	(0.333)	(0.424)	(0.304)	(0.326)
Price_Difference	-0.0164	0.110	0.0736	0.0987
	(0.879)	(0.324)	(0.457)	(0.366)
Market_LogRevenue	0.00348	0.00349	0.00367	0.00283
	(0.606)	(0.618)	(0.593)	(0.681)
Observations	3757	3757	3757	3757

Note: p-value in parentheses; $^{\star}p < 0.05$;$^{\star\star}p < 0.01$;$^{\star\star\star}p < 0.001$

Table EC.7 Robustness Check - Installer Level Dynamic Panels

Variables	(I) Market Transaction	(II) Market Transaction
Rating_Entropy_Mkt	1.501***	0.751*
	(0.000)	(0.016)
Rating_Entropy_Mkt^2	-2.456***	-1.609***
	(0.000)	(0.000)
Average_Rating_Mkt	-0.207	-0.141
	(0.169)	(0.476)
Average_Experience	0.0102	-0.00118
	(0.157)	(0.882)
Price_Difference_Mkt	0.0349	-0.313
	(0.865)	(0.089)
Market_LogRevenue	-0.0264	0.0403
	(0.502)	(0.051)
Market_Transaction$_t$		0.0238
		(0.685)
Market_Transaction$_{t-1}$		-0.00533
		(0.917)
Market_Transaction$_{t-2}$		0.198**
		(0.001)
Market Fixed Effect	Yes	No
Weighted State Dummies	Yes	Yes
Observations	642	421

Note: p-value in parentheses; $^{\star}p < 0.05$;$^{\star\star}p < 0.01$;$^{\star\star\star}p < 0.001$

Table EC.8 Robustness Check - Market Level Dynamic Panels

Variables	(I) Installer's Activity Level	(II) Installer's Activity Level	(III) Installer's Activity Level	(IV) Installer's Activity Level
Rating_Entropy_Others_1	0.512*			
	(0.013)			
Rating_Entropy_Others_2	-1.934***			
	(0.000)			
Rating_Entropy_Self_1		1.200***		
		(0.000)		
Rating_Entropy_Self_2		-3.072***		
		(0.000)		
Rating_Entropy_Others_1			0.834***	
			(0.001)	
Rating_Entropy_Others_2			-1.164*	
			(0.017)	
Rating_Entropy_Others_3			-1.763*	
			(0.020)	
Rating_Entropy_Self_1				1.746***
				(0.000)
Rating_Entropy_Self_2				-1.015
				(0.099)
Rating_Entropy_Self_3				-4.641**
				(0.008)
Observations	4562	4562	4562	4562
Adjusted R^2	0.631	0.633	0.632	0.633
AIC	13222.5	13204.4	13218.4	13203.5
BIC	13306.0	13287.9	13308.3	13293.4

Note: p-value in parentheses; $^{\star}p < 0.05$;$^{\star\star}p < 0.01$;$^{\star\star\star}p < 0.001$

Table EC.9 Alternative Inverted-U Testing: Spline Regressions (Installer Level)

Variables	(I) Market Transaction	(II) Market Transaction
Rating_Entropy_Mkt_1	1.197***	
	(0.000)	
Rating_Entropy_Mkt_2	-1.144**	
	(0.008)	
Rating_Entropy_Mkt_1		2.157***
		(0.000)
Rating_Entropy_Mkt_2		-2.100***
		(0.000)
Rating_Entropy_Mkt_3		0.356
		(0.606)
Observations	642	642
Adjusted R^2	0.720	0.732
AIC	1101.0	1074.3
BIC	1136.7	1114.4

Note: p-value in parentheses; $^{\star}p < 0.05$;$^{\star\star}p < 0.01$;$^{\star\star\star}p < 0.001$

Table EC.10 Alternative Inverted-U Testing: Spline Regressions (Market Level)

Variables	(I) Installer's Activity Level	(II) Installer's Activity Level	(III) Installer's Activity Level
Text_Dispersion_Self		11.42*	13.28*
		(0.032)	(0.016)
Text_Dispersion_Self^2		-34.86*	-37.87*
		(0.019)	(0.013)
Text_Dispersion_Other	14.22	13.10	14.26
	(0.086)	(0.113)	(0.080)
Text_Dispersion_Other^2	-59.45*	-54.83	-57.21*
	(0.035)	(0.051)	(0.039)
Average_Rating_Self	-0.436	-0.489	
	(0.086)	(0.062)	
Average_Rating_Others	-0.0164	-0.0166	
	(0.468)	(0.466)	
Average_Sentiment_Self			0.268
			(0.438)
Average_Sentiment_Others			0.0770
			(0.802)
Reviews_Count	0.0472*	0.0450*	0.0466*
	(0.000)	(0.000)	(0.000)
Experience	0.104	0.0946	0.100
	(0.124)	(0.164)	(0.144)
Price_Differences	-0.00831	0.0141	0.0249
	(0.936)	(0.893)	(0.814)
Market_LogRevenue	-0.0162*	-0.0158*	-0.0156*
	(0.002)	(0.002)	(0.002)
Constant	1.507*	0.747	0.236
	(0.029)	(0.348)	(0.782)
State Dummies	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Adjusted R^2	0.668	0.669	0.670
AIC	8853.4	8849.3	8843.7
BIC	8925.9	8933.9	8928.3

Note: p-value in parentheses; $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$
The text-based dispersion variables are derived by taking the mean, instead of median, of the pairwise cosine distances.

Table EC.11 Robustness Check - Installer Level Analysis with Text-based Dispersion from Mean Cosine Distances

	(I) Market Transaction	(II) Market Transaction	(III) Market Transaction
Rating_Entropy_Mkt			1.003*
			(0.023)
Rating_Entropy_Mkt^2			-1.079*
			(0.019)
Average_Rating_Mkt	-0.159		
	(0.252)		
Average_Sentiment_Mkt		-1.310**	-1.303**
		(0.008)	(0.008)
Average_Experience	0.0139	0.0108	0.00768
	(0.119)	(0.213)	(0.365)
Price_Difference_Mkt	0.388	0.411*	0.425*
	(0.055)	(0.035)	(0.043)
Market_LogRevenue	-0.0357	-0.0289	-0.0271
	(0.423)	(0.504)	(0.526)
Text_Dispersion_Mkt	20.58**	16.65*	14.67*
	(0.005)	(0.021)	(0.047)
Text_Dispersion_Mkt^2	-73.10**	-62.38**	-57.62*
	(0.002)	(0.008)	(0.016)
Constant	2.174	2.553	2.517
	(0.143)	(0.053)	(0.054)
Adjusted R^2	0.745	0.749	0.752
AIC	1077.6	1067.4	1064.6
BIC	1166.8	1156.6	1162.7

Note: p-value in parentheses; $^{*}p < 0.05$;$^{**}p < 0.01$;$^{***}p < 0.001$
The text-based dispersion variables are derived by taking the mean, instead of median, of the pairwise cosine distances.

Table EC.12 Robustness Check - Market Level Analysis with Text-based Dispersion from Mean Cosine Distances