# Do Noisy Customer Reviews Discourage Online Platform Sellers? Empirical and Text Analysis of a Solar Marketplace

**(Authors' names blinded for peer review)**

This paper

*Key words*: marketplace, reviews

## 1. Introduction

Solar energy is booming in the world. It is one of the fastest growing energy generation technology with a dazzling 34% growth worldwide in 2017 (Agency 2018). An important driver of this growth is increasing solar panel installations by electricity end-users. More and more, electricity end-users have been generating their own power with solar panels, reducing their reliance on electric utility companies. This type of solar generation skyrocketed in the last decade. For example, in the U.S., residential solar capacity increased by a factor of XY from 2012 to 2019 (US EIA XYZ). The annual residential solar panel installations are forecasted to grow 25% per year for the U.S (Weaver 2019, SEIA) with an even larger surge in the U.S after the passing of California Solar mandate (Pyper 2018).

There is an increasing trend of installing rooftop panels through online marketplaces. Consumer interest doubled in 11 states between 2017 to 2018, according to an analysis of website traffic (INC). An online solar marketplace is an innovative business model that eases the rooftop solar panel adoption process for electricity end-users. It essentially serves as an intermediary which connects buyers and installers of panels, making the process more transparent (Dorsey 2019).

In building an online marketplace, online reviews are considered as an essential functionality. In the literature, there are studies that investigate how the average customer *ratings* impact a single firm's sales. The consensus is that the average customer ratings can have significant impact

on sales, especially for products and services that entail searching and experiencing attributes (Zimmermann et al. 2018). In this paper, our primary goal is to empirically study the impact of *dispersion* of customer reviews on the performance metric of the platform, which is a composite of *many* firms. To the best of our knowledge, there is no prior work that has studied this topic.

Customer ratings are generally measured on a five-point scale. In this paper, we consider customer *reviews* that include both customer ratings and the review text made by verified buyers. Thus, our analysis employs recent text mining techniques as well as traditional statistical tools.

## 1.1. Relevant Literature

Contributions: 1) No prior study on the impact of reviews on the activity level of a firm (number of bids) in a marketplace. 2) The first that considers the impact of dispersion of competitor reviews on a firm's activity level on a platform. 3) No prior study on the impact of reviews on the number of matching in a platform where sellers have to take an action to acquire a potential customer. (I did an extensive search, and it is very surprising that there is no work about this! Just in case, I will continue to search more.)

Explanation of the literature and two additional papers for your reference: The vast majority of the literature on online ratings study the impact of average ratings on a firm's sales. Focusing on books or movies , several studies conclude that an improvement in the average rating of a product increases the product's sales (see, e.g., Chintagunta et al. (2010) and Chevalier and Mayzlin (2006)). Luca (2016) verifies that the average rating of a restaurant has a positive impact its revenue. There are also studies that show that the average ratings for a movie do not have a significant impact on its revenue (see, e.g., Duan et al. (2008))   The impact of the ratings dispersion is understudied in the literature, and there is no consensus about the impact. Clemons et al. (2006) find a positive correlation between the dispersion of ratings 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 variation in ratings 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. In the context of the movie industry, Zhang (2006) finds that variance of reviews does not play a significant role in movie openings.

Apart from these, Sun (2012) shows that a book's average rating interacts with the variance of its rating in its effect on book sales. Zimmermann et al. (2018) theoretically analyze a model to decompose the variance of product's ratings into ones caused by taste differences and quality differences and their impacts. (I did not include them on the slide as they would not clearly fit to one category - positive/negative/insignificant)

Our paper is also related to papers that investigates the effect of ratings on a single firm's performance metrics. In that stream, there is no consensus about the ultimate impact of dispersion

**Authors' names blinded for peer review**
Article submitted to *Manufacturing & Service Operations Management*; manuscript no.

3

of ratings on the firm's performance metric. Studies have demonstrated positive impacts (Chintagunta et al. 2010, Chevalier and Mayzlin 2006, Dellarocas et al. 2007), insignificant impact (Duan et al. 2008), and negative impacts in some instances (Wang et al. 2015).

In the literature, there are papers that show the positive impact of reviews on sales. There are also other papers that demonstrate (AVERAGE LIT (Literature considered average effect)).

Different from these papers, we took a perspective of the marketplace operator. The marketplace perspective is an important one, especially from the marketplace providers' perspectives. Many new businesses are running a marketplace business model, and have designed the customer ratings functionality an essential part of the platform experience (CITE SOMETHING). In our work, we use the total number of successful proposals on a relevant local market to gauge the health of the marketplace. Total number of success proposals as a performance metric is consistent with common business practices in the investment circle (Boris 2018, Galston 2017) as it is tied to a marketplace business's valuation.

Our objective is to understand the impact of review dispersion on the activity level of each participating supplier on the platform, which has not been studied before. Our study provides insights into the operation of a marketplace and ties reviews to marketplace matching metrics.

### 1.2. How Reviews Dispersion Impacts Activity Intensity(Literature Review)

In this section we describe several mechanisms by which reviews dispersion may impact installers activity intensity on a platform.

Previous studies have established the important of performance feedback on worker productivity. In a hospital setting Song et al. (2017) found a positive impact from public performance feedback to low-performing physicians. In a restaurant setting, coworker performances influence waiters own 'up-selling' behavior, a reflection of efforts, in an non-linear, inverse U-shape fashion.

The concept of **ratings dispersion** has been explored in marketing literature. For example, Luo et al. (2013) examined the brand ratings dispersion and its impact on firm values. In the economics literature, Marinovic (2015) modeled the phenomenon of performance feedback signal with a noise in a principal-agent model and illustrated feedback noise has potential of inducing agents efforts. Overall, the impact of feedbacks dispersion is less explored in an operations setting.

The impact of high ratings could be two-folded. On the one hand, high variations could be an indicator that the ratings scheme is functioning as it is designed - it rewards good installer and records the bad deeds of the bad ones. It could encourage installers to pursue more leads in order to get a chance to be evaluated. On the other hand, a high ratings variation could also be taken as

a sign of picky customers on the market. Installers fear of establishing bad permanent reputation will be more cautious when getting into a market of potentially picky customers.

In this study, we make use of the detailed installer level activities data. We explore not only the impact of ratings, but more importantly, the nuanced impact of the ratings dispersion and reviews variation.

### 1.3.   Overview

We first quantify the impact of dispersion with the activity intensity on an individual installer level; we then elevate our analysis to the platform level by connecting the impact of dispersion on local market level total transactions in relation to the dispersion in reviews.
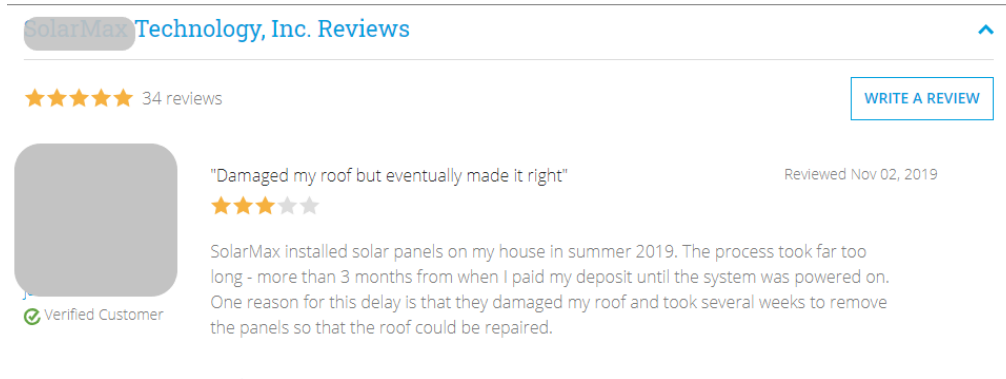
## 2.   Data and Setting

We analyze the interplay between customer reviews and firm activities (and outcomes) in an online marketplace for electricity end-users' solar panel installations. To do so, we collaborated with an online solar marketplace company, and obtained the full record of customer reviews and installer proposal activities on a monthly level from 2013 to 2018 in the marketplace. This data set is proprietary and it is the primary source of our analysis. We also complement the marketplace data with Tracking The Sun (TTS) data set from the Lawrence Berkeley National Laboratory. TTS is a comprehensive and publicly available data set on U.S. solar panel installations. Below, we provide details about our data and the setting of the online solar marketplace we study.

### 2.1.   Online Solar Marketplace

The solar marketplace (MKT) we study is an independent shopping website for electricity end-users (i.e., homeowners) who are interested in installing solar panels. The marketplace operates in 33 states of the U.S., and allows solar panel installers to maintain a profile, receive information on 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. Each installer provides service in a particular region. If the customer's location falls into an installer's service area, the marketplace informs the installer about the customer's arrival along with her information. Next, every informed installer decides whether to make a proposal to the customer. After the customer observes installer proposals, 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. Figure 1 provides an example of how customer reviews are displayed on the marketplace.

**SolarMax Technology, Inc. Reviews** ⌃

★★★★★ 34 reviews                                                                    **WRITE A REVIEW**

"Damaged my roof but eventually made it right"                                      Reviewed Nov 02, 2019
★★★☆☆

SolarMax installed solar panels on my house in summer 2019. The process took far too
long - more than 3 months from when I paid my deposit until the system was powered on.
One reason for this delay is that they damaged my roof and took several weeks to remove
the panels so that the roof could be repaired.

J...
✅ Verified Customer

**Figure 1**     **A sample customer review on the marketplace.**

|                       | Description                                                              |
|-----------------------|--------------------------------------------------------------------------|
| Installers            | 416 Unique Installers                                                    |
| Ratings and reviews   | 3607 pieces of review records with the rating, text content,timestamp    |
| Time span             | from 2013 to 2018                                                        |
| Monthly Records       | 6522 pieces                                                              |
| Supplementary         | Tracking the Sun data                                                   |

**Table 1     Main Data Source**

Note that the key decision of each installer in the marketplace is whether to make a proposal for each potential customer. In this context, we study how the dispersion of customer reviews impact the (i) *intensity of installer activity*, which is measured by how many proposals an installer makes per month, and (ii) number of matches, which is an important performance metric for the marketplace.

To answer these questions, we obtained rich panel data from the solar marketplace that contain all of its vetted installers across the U.S., installers' monthly activities and all customer reviews (text content and ratings) from the beginning of the marketplace (January of 2013), up to April 2018. Specifically, in our data set, we have observations about 416 installers about their monthly activities, i.e., the number of proposals made and the number of proposals won by each installer in every month, and 3607 pieces of customer reviews each with a rating, text content, time stamp, and the installer name with which the review is associated. Features of this data set are summarized in Table 1. We also collected the location information of each installer from their profiles, as illustrated in Figure 2.

It is perhaps worth mentioning that based on our conversions with the online marketplace, the marketplace actively reaches out to solar installers to recruit them to join to platform and help them set up their profile. So, unlike starting a physical business, installers' fixed cost of entry to the marketplace is negligibly small (if not zero). This is indeed the case in many different platforms (Haddad and Kleiner 2015). In light of this, in this study, we do not study the entry of installers
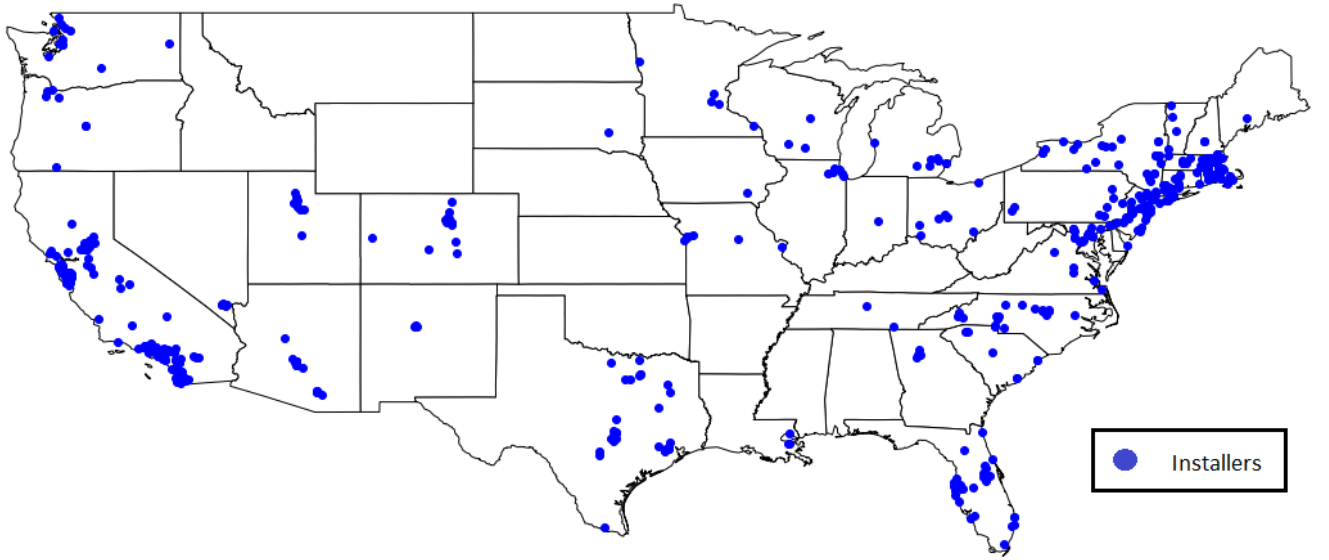
**Figure 2      Installers in our data set**

to the marketplace. Rather, we focus on installers' activity levels after they establish their profiles on the marketplace.

## 2.2.   Defining Local Market

Solar installation is a combination of product and service. As part of service, installers typically visit the customer site multiple times. Thus, each installer only operates within a certain geographical area. This means that installers compete "locally." That is, they only compete with installers that are relatively nearby. To capture this practical element, we identify what is called *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 state, county, or congressional district borders because it is common for installers to cross these artificial borders to serve customers. Instead, we use installer locations and the state-of-the-art advanced clustering algorithm called OPTICS (short for *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 the DBSCAN (Kanagala and Krishnaiah 2016). Among others, an important advantage of the OPTICS algorithm is that it does not require setting the number of clusters before running the algorithm as in $k$-means clustering; rather, it identifies the optimal number of clusters using the data. Because of these advantages, it has been applied in various contexts, ranging from political science (Davidson 2019) to geography (Teimouri et al. 2016).

## Installers Reviews
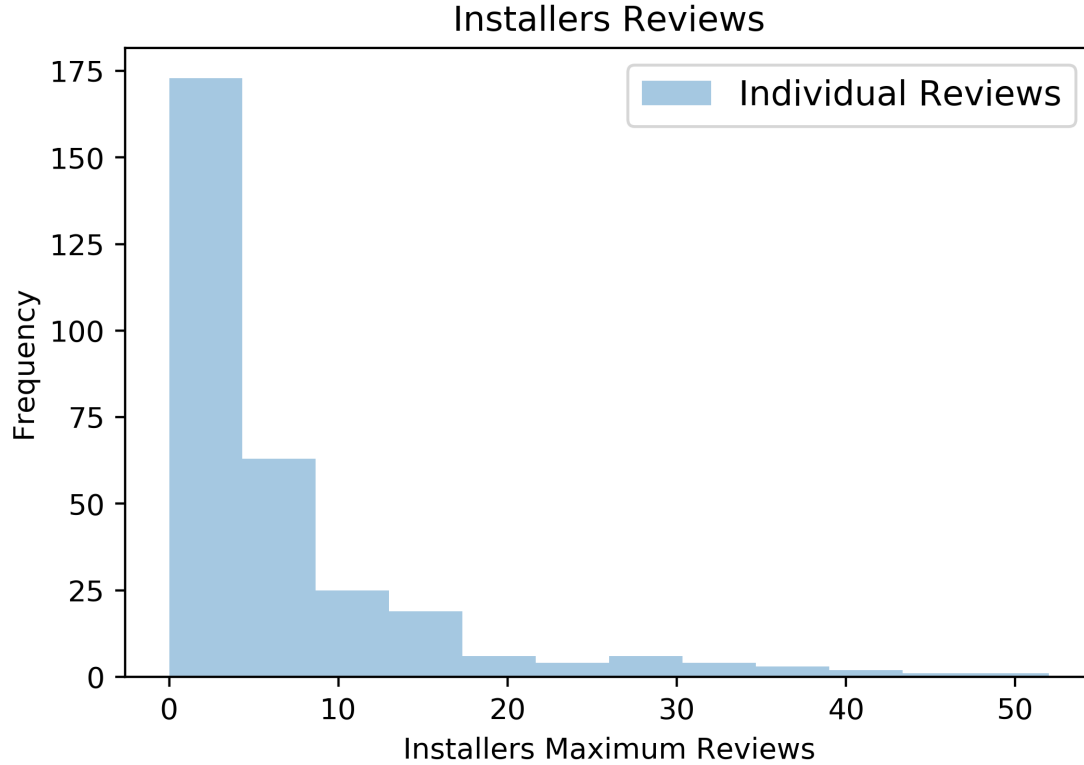


**Figure 3      Reviews Histogram by Installer**

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 2 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 the marketplace, we learned that the vast majority of customers get a quote from an installer within 100 miles of their property. We used 90 miles as the maximum distance input parameter to balance between having enough clusters to make use of the inherent variations and making sure that each cluster captures the local market condition. Based on this, the OPTICS algorithm generated 36 clusters.[1] Each of these clusters geographically defines a local market boundary. Figure 4 illustrates the centroid of each of these 36 clusters, which represents the centroid of each local market. Hereafter, for brevity, we refer to local markets simply as "*markets*."

---

[1] We also checked the robustness of our results by taking the maximum distance parameter as 100 miles in the OPTICS algorithm. Our insights remain to be valid in that alternative formulation.

8

Authors' names blinded for peer review
Article submitted to *Manufacturing & Service Operations Management*; manuscript no.

| State | Unique Installers | Total Number of Reviews | Total Number of Installations | Total Number of Quotes |
|-------|-------------------|-------------------------|-------------------------------|------------------------|
| CO | 13 | 799 | 168 | 13276 |
| MD | 10 | 895 | 144 | 4054 |
| WA | 9 | 902 | 35 | 1266 |
| TX | 27 | 987 | 90 | 12557 |
| FL | 21 | 994 | 141 | 9047 |
| CT | 10 | 1037 | 78 | 2746 |
| NC | 16 | 1100 | 95 | 7066 |
| NJ | 26 | 1674 | 223 | 8215 |
| NY | 32 | 2790 | 265 | 15128 |
| MA | 36 | 3519 | 507 | 19028 |
| CA | 95 | 7703 | 1472 | 98597 |

**Table 2     Top 10 States**
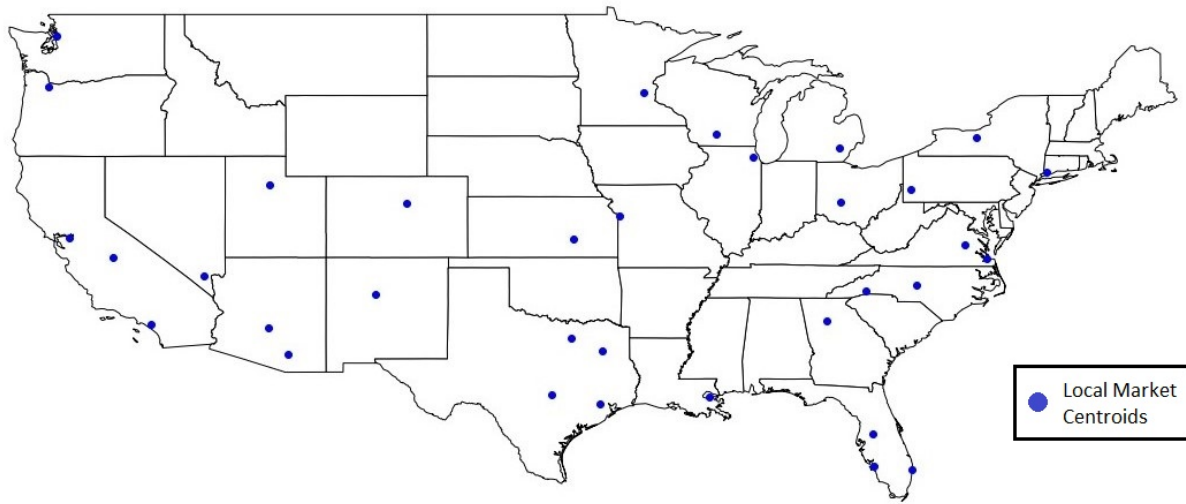


**Figure 4     Local Market Centroids**

## 2.3.    Measuring the Dispersion in Customer Reviews

In our analysis, a key explanatory variable is the dispersion in reviews, which can be measured with numeric ratings (from 1 to 5 stars) or the text content. Below, we first describe how we measure the dispersion of reviews based on ratings. We then explain the innovative word embedding model we use to measure the dispersion of reviews based on the text data. Our base empirical analysis uses the rating-based dispersion as a variable. We then check the robustness of our results by adding the text-based review dispersion as a separate variable in our analysis.

### 2.3.1.    Measuring Rating-Based Dispersion

We measure the rating-based dispersion by calculating the *entropy* of ratings. In information theory, the entropy is a common way to measure the uncertainty in a random variable's possible realizations. In our setting, because the marketplace

has a 5-star rating system, the entropy of ratings is

$$H(R) = -\sum_{i=1}^{5} \text{Prob}(\text{Rating} = i) \ln(1/\text{Prob}(\text{Rating} = i)). \tag{1}$$

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 Rating_Entropy_Self$_{i,t}$ that represents the entropy of each installer $i$'s own reviews. This is calculated on the set of reviews that are associated with installer $i$ up to month $t$. Recalling that market boundaries are defined as in Section 2.2, the second variable is Rating_Entropy_Others$_{i,t}$ that is the rating entropy of all other installers in installer $i$'s market, up to month $t$. In one of our extension sections, we will also consider the rating entropy on the market level. Thus, our third variable is Rating_Entropy_Mkt$_{m,t}$ that represents the entropy of all ratings in the local market $m$, up to that month $t$. Again, the market is defined as in Section 2.2.

We choose to use entropy to measure the dispersion of ratings over other more common measures, such as variance or coefficient of variance, for a few practical reasons: 1) The variance has a higher than 0.7 correlation with the average rating. Entropy measures exhibit a lower level of correlation with average ratings. That enabled us to include both variables in the model. 2) entropy measure provides a higher precision (for our data): two installers with very small ratings variance differences tend to show a larger entropy measure difference.

### 2.3.2. Measuring Text-Based Dispersion via the Language Model BERT
In our analysis, we leverage the rich information in text reviews. To do so, we use the state-of-the-art deep learning based language model called BERT (short for *Bidirectional Encoder Representations from Transformers*). BERT was developed by Google, and published in 2018 (Devlin et al. 2018). Since then, it has been widely used in practice. For example, Google Search Engine has incorporated this method in October 2019 (REFERENCE). To the best of our knowledge, there is yet no other paper in operations management literature that considers this word embedding technique. We will provide more details about this technique toward the end of this section.

We apply the following steps to measure the text-based dispersion. First, we use the BERT model to convert every piece of review text to a numerical vector. Second, we normalize each vector to unit length. Third, we measure the cosine similarity to find the similarity between every two review

Authors' names blinded for peer review

10                  Article submitted to *Manufacturing & Service Operations Management*; manuscript no.

vectors. We note that normalization and the cosine similarity measure are standard to identify the similarity between two vectors. See, e.g., Hoberg and Phillips (2016). 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. This angle represents the similarity in the orientation of two vectors. If the angle is 0, the two vectors are at the same orientation and hence the similarity is 1, which is maximum. After this step, we identify the cosine distance between every two review vectors from the fact that the cosine distance between two normalized vectors $V_1$ and $V_2$ equals 1 minus the cosine similarity between the two. This distance reflects how different two reviews are from each other. As the final step, we calculate the dispersion in sets of text reviews, *text-based dispersion*, in short, by enumerating all pairwise distances of reviews in that set and taking their statistical median (the 50th percentile). For example, for a set of 10 text reviews, we have 45 $(=\binom{10}{2})$ pairwise distances. Finding the text-based dispersion for this set requires computing the median distances 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.

As a result of this procedure, similar to the ratings entropy, we create the following three variables, each measuring the text-based dispersion in a different dimension: (i)$Text\_Dispersion\_Self_{i,t}$: Dispersion in installer $i$'s own text reviews up to month $t$. (ii) $Text\_Dispersion\_Others_{i,t}$: Dispersion in the text reviews of all other installers in the installer $i$'s local market up to month $t$. (iii)$Text\_Dispersion\_Mkt_{m,t}$: Dispersion in text reviews of all installers in market $m$ up to month $t$.

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 called 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. Specifically, the BERT model has two distinct advantages. 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 meaning. 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.

## 3. Installer-Level Analysis & Results

This section answers the following two questions: (i) How does the dispersion in an installer's reviews impact the number of proposals generated by the installer, which is referred to as the installer's *activity level*. (ii) How does the dispersion in competitor installers' reviews impact the installer's activity level?

To answer these questions, we run a regression model where the response variable is $Installer\_Activity_{i,m,t}$ and calculated as log(proposals generated by installer $i + 1$) in the market $m$ during month $t$. In our regression, first, we only use numerical ratings. Later, as a separate analysis, we will use text-based dispersion variables in our regression.

### 3.1. Model & Controls

Recalling the entropy variables defined in Section 2.3.1, we run the following regression model on our panel data:

$$
\begin{aligned}
Installer\_Activity_{i,m,t+1} = & \beta_0 + \beta_1 Rating\_Entropy\_Self_{i,t} + \beta_2 Rating\_Entropy\_Self_{i,t}^2 \\
& + \beta_3 Rating\_Entropy\_Others_{i,t} + \beta_4 Rating\_Entropy\_Others_{i,t}^2 \\
& + Controls_{i,m,t} + \alpha_i + \epsilon_{i,t+1}.
\end{aligned} \tag{2}
$$

Here, $\epsilon$ 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 $\alpha_i$ 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 the random-effect model is preferred while the alternative is the fixed-effect model. With a p-value $< 0.0001$,

|                        | Obs  | Mean     | SD       | Min       | Max      |
|------------------------|------|----------|----------|-----------|----------|
| Ratings Average        | 4521 | 4.524005 | 1.329408 | 1         | 5        |
| Reviews Count          | 4521 | 5.211429 | 6.650352 | 0         | 52       |
| Entropy Own Reviews    | 4521 | .0937851 | .2147703 | 0         | 1.209574 |
| Entropy Others Reviews | 4521 | .2098061 | .1864398 | 0         | 1.070593 |
| Experience             | 4521 | 5.693455 | 5.748559 | 0         | 43       |
| Price Diff             | 4521 | -.033842 | .3810107 | -2.171179 | 6.641788 |
| Market Rev (Log)       | 4521 | 11.77942 | 8.023967 | 0         | 22.30267 |

**Table 3          Summary Statistics Individual Level**

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. Thus, we focus on (2) with the installer-level fixed effect $\alpha_i$ that controls for time-invariant characteristics of each installer.

The regression (2) includes various additional installer-level or market-level control variables (Controls$_{i,m,t}$). To account for the state-level renewable policy effects, we include state dummies denoted by "State" as a control variable. We have 33 such variables. We account for the impact of the solar panel prices on installers' activities by considering Price_Difference$_{i,t}$ as another control variable. We use the TTS data to find each installer's price for 1 KW solar panel via matching name and zipcode. In practice, price per KW is a common way to assess the price of a solar panel as solar systems vary in size. Based on this, we compute the variable Price_Difference$_{i,t}$ as 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 self average rating of each installer $i$ Avg_Rating_Self$_{i,t}$ as well as the average ratings of its competitors Avg_Rating_Others$_{i,t}$ in the market for month $t$. We also control for the experience of the installer by considering the variable Experience$_{i,t}$ that is the logarithm of number of years the installer has been installing solar systems up to month $t$. We obtain this information from each installer's website. Another control variable in (2) is Market_Revenue$_{m,t}$ that measures the 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 2.2 with the TTS data to capture total solar installations opportunities in the market. Finally, we consider the count of each installer $i$'s self reviews up to (and including) month $t$, and denote it by Review_Counts$_{i,t}$.

Tables 3 and 4 below present the summary statistics and the correlation matrix. By Table 4, correlations are in the expected direction and do not hurt the validity of regression analysis.

### 3.2. Results

Table 5 presents results estimated by three panel regression models based on (2). Columns (1) through (3) of Table 5 correspond to results obtained by using different set of explanatory variables in the regression. The results in the column (3) correspond to the ones estimated by (2); others

| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| (1) Entropy (Others) | 1.000 | | | | | | | |
| (2) Entropy (Self) | -0.075 | 1.000 | | | | | | |
| (3) Own Average Ratings | 0.063 | -0.088 | 1.000 | | | | | |
| (4) Reviews Count | 0.052 | 0.252 | 0.206 | 1.000 | | | | |
| (5) Others' Average Ratings | -0.514 | 0.022 | -0.012 | -0.040 | 1.000 | | | |
| (6) Experience (Log) | 0.053 | -0.009 | 0.024 | 0.130 | -0.092 | 1.000 | | |
| (7) Price (Diff) | -0.034 | -0.008 | 0.004 | -0.029 | 0.016 | -0.033 | 1.000 | |
| (8) Local Market Revenue (Log) | -0.102 | -0.046 | -0.029 | -0.064 | 0.040 | 0.525 | -0.064 | 1.000 |

**Table 4      Correlation Individual Level**

| | (1) | (2) | (3) |
|---|---|---|---|
| | F.Activity | F.Activity | F.Activity |
| Average Rating | -0.830* | -0.788* | -0.773* |
| | (0.000) | (0.000) | (0.000) |
| Average Ratings(Others) | -0.00976 | -0.00613 | -0.00491 |
| | (0.644) | (0.790) | (0.831) |
| Reviews Count | 0.0503* | 0.0473* | 0.0420* |
| | (0.001) | (0.000) | (0.000) |
| Experience | 0.220* | 0.213* | 0.213* |
| | (0.001) | (0.002) | (0.000) |
| Price Diff | 0.0785 | 0.0798 | 0.0856 |
| | (0.351) | (0.343) | (0.309) |
| Market Revenue | -0.0172* | -0.0168* | -0.0160* |
| | (0.000) | (0.000) | (0.001) |
| Entropy Others | | 1.759* | 1.754* |
| | | (0.000) | (0.000) |
| Entropy Others$^2$ | | -2.615* | -2.686* |
| | | (0.000) | (0.000) |
| Entropy(Self) | | | 1.743* |
| | | | (0.000) |
| Entropy(Self)$^2$ | | | -3.356* |
| Constant | 2.657* | 2.481* | 2.664* |
| | (0.000) | (0.000) | (0.001) |
| Fixed Effect | Yes | Yes | Yes |
| State Dummies | Yes | Yes | Yes |
| Observations | 4200 | 4200 | 4200 |
| r2 | 0.640 | 0.644 | 0.646 |
| aic | 12189.1 | 12156.3 | 12131.6 |
| bic | 12246.1 | 12226.0 | 12214.1 |

**Table 5      Individual Level Analysis**

are estimated by considering only some of these variables in the regression. In particular, column (1) corresponds to results when entropy variables are not considered at all whereas column (2) are the estimates when we include only others' entropy ( Rating_Entropy_Others_$i, t$).

Our estimates in Table 5 identify three key results. First, the set of variables representing "noise" or dispersion of ratings have a significant impact on the activity level of an installer in the marketplace. Second, the entropy of an installer's own ratings has a positive and statistically significant effect on its activity level (because $\beta_1$ in equation (2) is found to be positive and significant in column(3) of Table 5), and the second-order effect ($\beta_2$ in equation (2))) of an installer's ratings entropy is negative and statistically significant. Combining these two effects, the regression estimates indicate that the dispersion of an installer's own ratings increases the installer's activity level if and only if the aforementioned dispersion is small. When the dispersion of its own ratings is large, any additional dispersion in the installer's own ratings deters the its activity in the marketplace.

Third and most important, our estimation shows that the entropy of competitor ratings impacts an installer's activity level in the *same* way as the entropy of the installer's own ratings. That is, the dispersion of competitor ratings increases the installer's activity level if and only if the aforementioned dispersion is small. When the dispersion of competitor ratings is large, any additional dispersion in competitor ratings deters the installer's activity in the marketplace.

Figures 6 and 5 illustrate the explained non-linear effects of the rating entropy on installer's activity level. To generate the marginal effects displayed in Figures 6 and 5, we use the estimated regression coefficients in the column (3) of Table 5. As is apparent from these figures, the installer's activity level first increases then decreases with the rating dispersion, and that is true regardless of whether installer's own ratings or its competitors' ratings are considered.

Finally, it might be worth noting that in the first two columns average of own or others ratings did not appear to be significantly linked with activity intensity levels. Yet after we controlled for dispersion, we noticed a negative effect associated with installers' own average. Installers appear to extend fewer proposals as their own ratings increase, after all other factors are controlled for. We think this is a phenomenon that installers are more selective after they attained a high rating on the platform, perhaps acting out of confidence that with a higher rating, their proposals are more likely to be accepted by customers.

## 4. Market-Level Analysis & Results

An important performance metric for the marketplace is the number of matches (i.e., agreed proposals) between installers and customers on the marketplace. This section estimates how the dispersion of ratings in a market impacts the number of matches in that market.

For this estimation, we use the regression equation in (3). Our data include the number of agreed proposals for each installer $i$ in each month $t$. To create our dependent variable 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 logarithmic transformation of that sum. Formally,
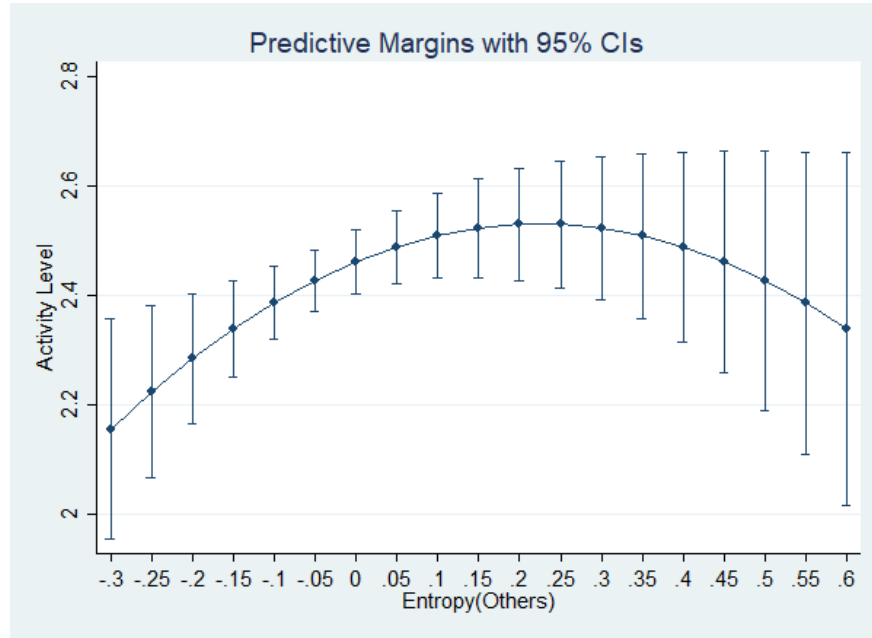
**Figure 5** **Marginal Impact of the Entropy of Other Installers' Ratings (in the same market) on the Installer's Activity Level**
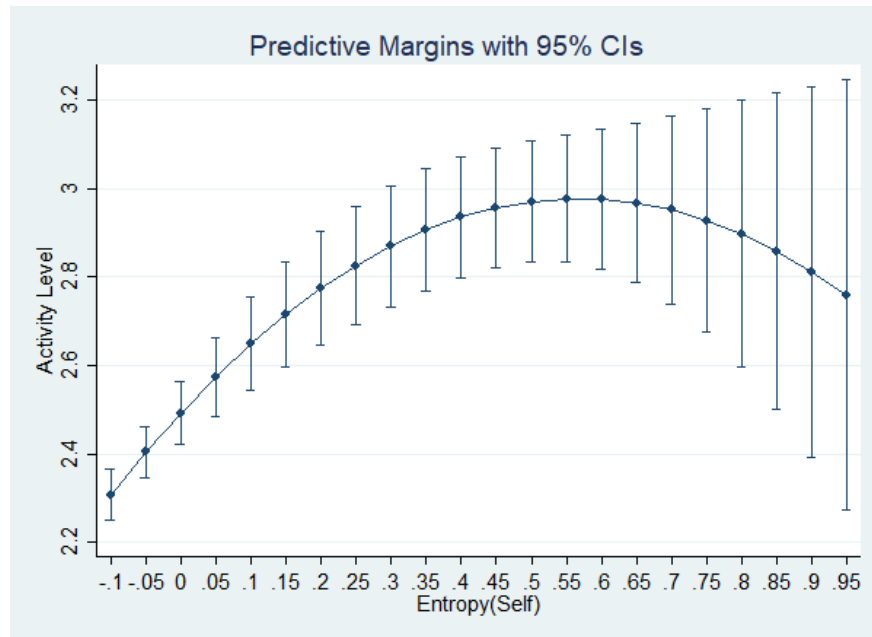


**Figure 6** **Marginal Impact of the Entropy of the Installer's Own Ratings on its Activity Level**

$Market\_Transaction_{m,t+1} = \ln\left(\sum_{i\in\text{Market } m}\text{Successful\_Proposals}_i + 1\right)$. With this, we convert the installer-level monthly panel data to market-level monthly panel data so that we can exploit the market-level rating dispersion to identify its impact on the local market outcome.

$$Market\_Transaction_{m,t+1} = \beta_5 + \beta_6 Rating\_Entropy\_Mkt_{m,t} + \beta_7 Rating\_Entropy\_Mkt^2_{m,t}$$

| VARIABLES | N | mean | p50 | min | max |
|---|---|---|---|---|---|
| Entropy | 611 | 0.201 | 0.169 | 0 | 1.055 |
| Text-based Entropy | 611 | 0.131 | 0.125 | 0.0689 | 0.366 |
| Mkt Revenue | 611 | 8.029 | 10.50 | 0 | 22.30 |
| Rating Avg | 611 | 4.866 | 4.988 | 3 | 5 |
| Sentiment Score | 611 | 0.672 | 0.689 | 0.0593 | 0.866 |
| Experience(Avg) | 611 | 5.922 | 6.375 | 0 | 27 |
| Price Diff(Avg) | 611 | -0.0160 | 0 | -0.504 | 1.312 |
| Transaction | 611 | 1.019 | 0.693 | 0 | 4.522 |

**Table 6    Market Level Summary Statistics**

$$+ Controls_{m,t} + \epsilon_{m,t+1}. \tag{3}$$

Here, $Rating\_Entropy\_Mkt_{m,t}$ is as defined in Section 2.3.1 and equals the entropy of ratings up to (and including) month $t$ from all installers in the market $m$.

We also use various control variables denoted by $Controls_{m,t}$. We control for the state of the market. We created 33 state dummies to represent 33 different state represented in the data set. Some market span across more than one state. In that case, we calculated state representation fractions and assign its value (between 0 to 1) to the state indicator variable ( i.e. market 1 has 25% of installers from state CA and 75% from NV, then we assign 0.25 to dummy STATE_CA and 0.75 to dummy STATE_NV). Similar to the installer-level analysis in Section 3, we created the $Avg\_Experience_{m,t}$ variable to represent the average experience of installers in the market $m$ up to and including month $t$. We calculated this by averaging installers' experiences level across the local market $m$. In parallel to the installer-level analysis, we use the $Avg\_Rating\_Mkt_{m,t}$ to control for the average rating of all installers in the market $m$ until (and including) month $t$. As another control, we created $Review\_Count\_Mkt_{m,t}$ that measures the total number of reviews by all installers in market $m$ up to and including month $t$. We also control for the difference of average unit price between marketplace and off-marketplace. We use $Price\_Diff\_Mkt_{m,t}$ to denote this variable. Finally, we use $Market\_Revenue_{m,t}$ as a control where it is as defined in Section 3.Summary statistics can be found in Table 6; the correlation coefficients among variables are presented in Table 7.

## 4.1.    Results

The results are presented in Table 4.1. Column (1) are the estimates obtained by running the regression in (3) in the absence of $Rating\_Entropy\_Mkt_{m,t}$ and $Rating\_Entropy\_Mkt^2_{m,t}$ variables, and Column (2) are the estimates obtained by running the regression in (3) considering all variables in (3).

he "noise" portion of the ratings remains a significant factor. The estimate suggest that on the market level reviews dispersion is directly linked to higher number of total matches (proposals

| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| (1) Transaction | 1.000 | | | | | | |
| (2) Activity | 0.748 | 1.000 | | | | | |
| (3) Entropy | 0.226 | 0.380 | 1.000 | | | | |
| (4) Rating Avg | -0.026 | -0.151 | -0.601 | 1.000 | | | |
| (5) Mkt Revenue | 0.332 | 0.282 | 0.146 | -0.036 | 1.000 | | |
| (6) Experience(Avg) | 0.210 | 0.270 | 0.215 | -0.111 | 0.484 | 1.000 | |
| (7) Price Diff(Avg) | -0.116 | -0.102 | -0.062 | 0.000 | -0.045 | 0.012 | 1.000 |

**Table 7    Correlation Market Level**

| | (1) | (2) |
|---|---|---|
| | F.Transaction | F.Transaction |
| Rating Avg | -0.273 | -0.185 |
| | (0.071) | (0.212) |
| Mkt Revenue | -0.0846 | -0.0663 |
| | (0.057) | (0.123) |
| Experience(Avg) | 0.0197* | 0.0138 |
| | (0.013) | (0.073) |
| Price Diff(Avg) | 0.239 | 0.287 |
| | (0.282) | (0.194) |
| Entropy | | 1.060*** |
| | | (0.000) |
| Entropy$^2$ | | -1.610*** |
| | | (0.000) |
| Constant | 3.443** | 2.909* |
| | (0.009) | (0.025) |
| Observations | 642 | 642 |
| Market Fixed Effects | Yes | Yes |
| Weighted State Dummies | Yes | Yes |
| r2 | 0.739 | 0.747 |
| aic | 1075.6 | 1059.6 |
| bic | 1156.0 | 1148.9 |

**Table 8    Market Level Analysis**

accepted), as reflected in the positive and statistically significant coefficient ($\beta_5 = 1.628, p < 0.01$). We also note that the second order effect is negative as the coefficient estimates associated with the square term is negative ($\beta_6 = -1.160, p < 0.01$) and statistically significant. Combining these two effects, the regression estimates indicate that the dispersion of installers ratings on a local market increases the local market matches if and only if the aforementioned dispersion is small. When the dispersion of market level ratings is large, any additional dispersion in the market level dampens the market level matching activity.

We further illustrate this point with a margins plot using coefficients generated from estimates in column (2) of table 4.1 in figure 7. As it is apparent from this figure, the market level matching

Authors' names blinded for peer review

18          Article submitted to *Manufacturing & Service Operations Management*; manuscript no.
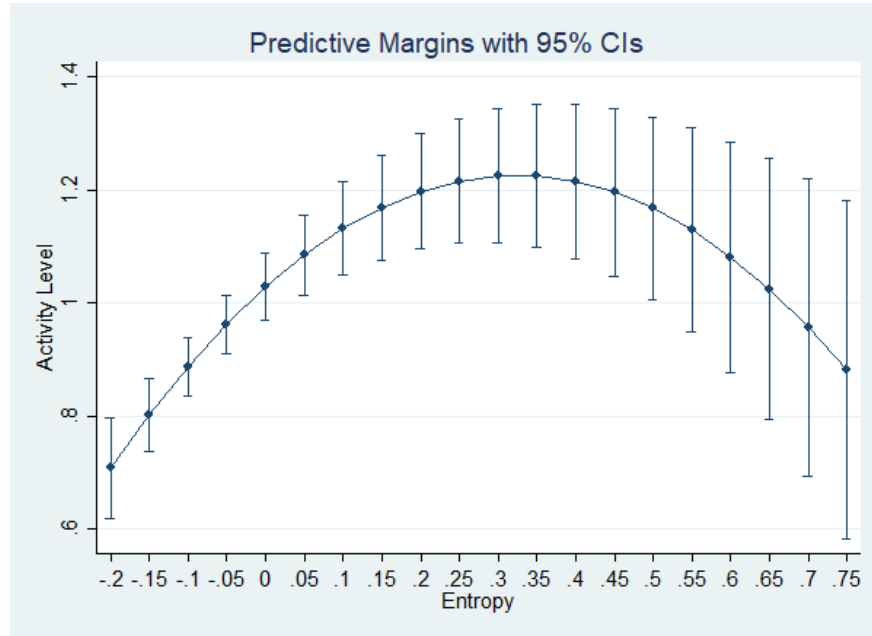


**Figure 7**     **Marginal Impact of Market Reviews Entropy of Reviews on Market Level Matching**

increase then decreases with the ratings dispersion.

Finally, the results on the local market level also suggest that after controlling for local market conditions, installer experience, price, and state, the average rating is not significantly associated with the market level performance.

## 5. Text Mining

In this section, we incorporate various methods to leverage the rich text information in reviews. We first use NLP method to generate sentiment score of each reviews. We also apply BERT model to perform word embedding, and generated measures for texts dispersion. We replace quantitative metrics derived from ratings with text mining measures in our analysis.

### 5.1. Sentiment Score of reviews Texts

In this section, we introduce methods to generate sentiment scores on the reviews texts. We use VADER model to generate text sentiment score. VADER, developed by Hutto and Gilbert (2014) as a 'Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text'. Since reviews texts shares many stylistic similarities with social media text, this is an appropriate approach. For a piece of text, the language model produces a compound sentiment score from -1 to 1, with 1 representing very positive and -1 very negative sentiments.

Example 1:

| | (1) | (2) | (3) |
|---|---|---|---|
| | F.Activity | F.Activity | F.Activity |
| Average Rating | -0.956* | | |
| | (0.000) | | |
| Average Ratings(Others) | -0.00337 | | |
| | (0.870) | | |
| Avg Sentiment Score | | -0.416 | -0.371 |
| | | (0.106) | (0.172) |
| Avg Sent(Others) | | -0.167 | -0.181 |
| | | (0.413) | (0.369) |
| Reviews Count | 0.0428* | 0.0450* | 0.0392* |
| | (0.000) | (0.000) | (0.000) |
| Experience | 0.207* | 0.190* | 0.196* |
| | (0.001) | (0.001) | (0.000) |
| Price Diff | 0.0931 | 0.100 | 0.120 |
| | (0.302) | (0.259) | (0.179) |
| Market Revenue | -0.0165* | -0.0171* | -0.0164* |
| | (0.001) | (0.000) | (0.001) |
| Text-based Entropy(Self) | 4.338* | 4.489* | 3.960* |
| | (0.000) | (0.000) | (0.000) |
| Text-based Entropy(Self)$^2$ | -21.40* | -19.92* | -16.42* |
| | (0.000) | (0.000) | (0.000) |
| Text-based Entropy(Others) | -0.790 | -0.728 | -1.480 |
| | (0.628) | (0.648) | (0.356) |
| Text-based Entropy(Others)$^2$ | -40.07* | -20.20 | -6.617 |
| | (0.019) | (0.319) | (0.737) |
| Entropy(Self) | | | 1.601* |
| | | | (0.000) |
| Entropy(Self)$^2$ | | | -3.279* |
| | | | (0.000) |
| Entropy(Others) | | | 0.303 |
| | | | (0.123) |
| Entropy(Others)$^2$ | | | -2.357* |
| | | | (0.000) |
| Constant | 2.370* | 2.085* | 2.323* |
| | (0.000) | (0.000) | (0.000) |
| Observations | 4109 | 4192 | 4192 |
| r2 | 0.646 | 0.644 | 0.649 |
| aic | 11880.2 | 12094.8 | 12049.4 |
| bic | 11968.7 | 12183.5 | 12163.6 |

**Table 9      Individual Level Analysis with Variables Derived from Text Analysis**

*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.*

**Authors' names blinded for peer review**

20          Article submitted to *Manufacturing & Service Operations Management*; manuscript no.

|                          | (1)           | (2)           | (3)           |
| ------------------------ | ------------- | ------------- | ------------- |
|                          | F.Transaction | F.Transaction | F.Transaction |
| Rating Avg               | -0.262+       |               |               |
|                          | (0.095)       |               |               |
| Sentiment Score          |               | -1.322*       | -1.215*       |
|                          |               | (0.006)       | (0.011)       |
| Experience(Avg)          | 0.0127+       | 0.00856       | 0.00529       |
|                          | (0.082)       | (0.230)       | (0.458)       |
| Price Diff(Avg)          | 0.0877        | 0.127         | 0.127         |
|                          | (0.746)       | (0.640)       | (0.619)       |
| Mkt Revenue              | -0.0458       | -0.0296       | -0.0216       |
|                          | (0.253)       | (0.448)       | (0.574)       |
| Text-based Entropy       | 7.394+        | 5.963         | 3.237         |
|                          | (0.082)       | (0.152)       | (0.429)       |
| Text-based Entropy$^2$   | -19.33+       | -16.31        | -10.23        |
|                          | (0.066)       | (0.116)       | (0.314)       |
| Entropy                  |               |               | 1.376*        |
|                          |               |               | (0.000)       |
| Entropy $^2$             |               |               | -1.379*       |
|                          |               |               | (0.001)       |
| Constant                 | 2.900*        | 2.573*        | 2.424*        |
|                          | (0.045)       | (0.038)       | (0.046)       |
| Observations             | 611           | 611           | 611           |
| r2                       | 0.752         | 0.755         | 0.760         |
| aic                      | 1010.3        | 1001.8        | 992.3         |
| bic                      | 1098.6        | 1090.1        | 1089.4        |

**Table 10      Market Level Analysis with Variables Derived from Text Analysis**

The above review received a compound sentiment score of 0.8622 and a five-star rating.

Example 2:

*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 Esposito is also known as (...).*

This received -0.7184 sentiment score and an one-star rating.

## 5.2.  Analysis using Variables Derived From Text Mining

We move on to discuss the analysis which we conduct with variables derived from text mining. We have introduced two models BERT and VADER which are used for construct dispersion and average sentiment measures, respectively. Overall, the sentiment score correlates with ratings significantly(corr = 0.8239), while the text-based dispersion measure correlates with the entropy measures on a lower level (corr = 0.4309). The correlation between sentiment score and ratings

is further illustrated in figure 8 with a violinplot. And for the relationship between text-based dispersion and ratings entropy, we present a scatter plot with a regression line in figure 9.

We are interested to determine if installers and market level activities can also be attributed to sentiment scores and text-based dispersion measures. We use variables $\text{Avg\_Sent\_Self}_{i,t}$ in place of $\text{Avg\_Rating\_Self}_{i,t}$, $\text{Avg\_Sent\_Others}_{i,m,t}$ in place of $\text{Avg\_Others}_{i,m,t}$ to represent individual ratings in individual level analysis, $\text{Avg\_Sent\_Mkt}_{m,t}$ in place of $\text{Avg\_Mkt}_{m,t}$ in market level analysis, and ran the same regression models. The results are presented in table 9 for installer level analysis and table 10 for market level analysis.

In table 9, column (1) through (4) contains results obtained by using different set of explanatory variables. Specifically, (1) and (2) contains results including only Text_Based_Dispersion_Others, where column (1) model uses average ratings and column (2) model uses average sentiment score. Moving on, the models used in column (3) and (4) have a parallel structure, but included installers' own text-based dispersion.

We observe the same type of inverse U shape for the marginal impact of text-based dispersion.

$$
\begin{aligned}
\text{Installer\_Activity}_{i,m,t+1} =& \beta_0 + \beta_1 \text{Text\_Dispersion\_Self}_{i,t} + \beta_2 \text{Text\_Dispersion\_Self}^2_{i,t} \\
& + \beta_3 \text{Text\_Dispersion\_Others}_{i,t} + \beta_4 \text{Text\_Dispersion\_Others}^2_{i,t} \\
& + \text{Controls}_{i,m,t} + \alpha_i + \epsilon_{i,t+1}.
\end{aligned} \tag{4}
$$

$$
\text{Market\_Transaction}_{m,t+1} = \beta_5 + \beta_6 \text{Text\_Dispersion\_Mkt}_{m,t} + \tag{5}
$$
$$
\beta_7 \text{Text\_Dispersion\_Mkt}^2_{m,t} + Controls_{m,t} + \epsilon_{m,t+1}.
$$

## 6.    Robustness Check
### 6.1.    Endogeneity

We devised several empirical strategies to mitigate the potential drawbacks of endogeneity or omitted variables in our analysis. Regarding the individual level analysis, endogeneity could occur if there are unobserved factors that is significantly correlated with ratings dispersion that is also correlated with the activity intensities.

Consider that we omitted a variable that captures installer professionalism or motivation, which we denote as $pro_{i,t}$. The actual function should be

**Figure 8      Sentiment Scores and Ratings**
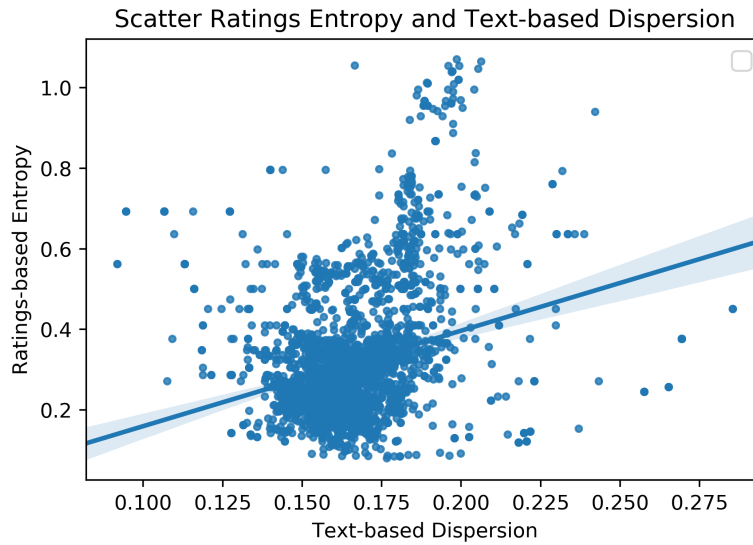


**Figure 9      Text-based Dispersion and Entropy of Ratings**

$$\begin{aligned}
\text{Installer\_Activity}_{i,m,t+1} = {} & \beta_0 + \delta \text{pro}_i t + \beta_1 \text{Rating\_Entropy\_Self}_{i,t} + \beta_2 \text{Rating\_Entropy\_Self}_{i,t}^2 \\
& + \beta_3 \text{Rating\_Entropy\_Others}_{i,t} + \beta_4 \text{Rating\_Entropy\_Others}_{i,t}^2 \\
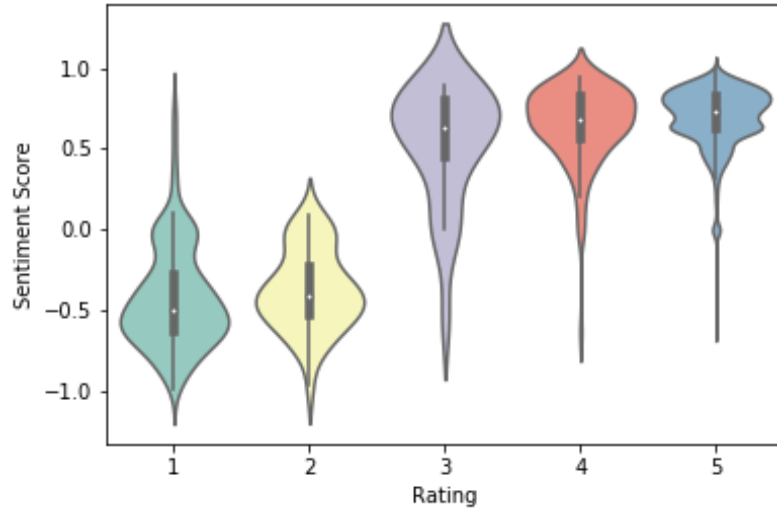& + \text{Controls}_{i,m,t} + \alpha_i + \epsilon_{i,t+1}.
\end{aligned} \tag{6}$$

We argue that $pro_{it}$ would be *negatively* correlated with reviews dispersion – professional installers would be more motivated than others to deliver consistent products and services (CITE

some thing).

In this case, the presence of omitted variable deflated the estimates of $\beta$ (CITE ECONOMET-RIC).

## 6.2.   Robustness with different local market division

Although many similar studies used ZIP code to difine local markets(cite something from IO), we used unsupervised algorithm (OPTICS) to determine the market grouping. OPTICS algorithm requires a few parameter inputs: X, Y and Z. We used parameter XX after performing grid-search on a parameter space XXX and use Calinski-Harabasz Index to assess the appropriateness of the clustering.

In addition, we used 4 digit ZIP code to define a market and the results are consistent ( INSERT RESULTS); we also use other OPTICS parameter and the results are consistent.

## 6.3.   Dynamic Panel model

In our main analysis we include fixed effect for each installer to account for time invariant factors. We use a dynamic panel model to perform robustness check. The inclusion of lagged dependent variables ( Activity Intensity) aim to control for unobserved heterogeneity that may influence changes in the dependent variable and is time variant. For individual level estimation, the equation we estimate is changed into the following:

$$\text{Installer\_Activity}_{i,m,t+1} = \beta_0 + \gamma_0 \text{Installer\_Activity}_{i,t} + \gamma_1 \text{Installer\_Activity}_{i,t-1} +$$

$$\beta_1 \text{Rating\_Entropy\_Self}_{i,t} + \beta_2 \text{Rating\_Entropy\_Self}_{i,t}^2 + \beta_3 \text{Rating\_Entropy\_Others}_{i,t} + \beta_4 \text{Rating\_Entropy\_Others}_{i,t}^2$$

$$+ \text{Controls}_{i,m,t} + \epsilon_{i,t+1}. \tag{7}$$

The inclusion of the lagged dependent variable is to try to control for unobserved heterogeneity that may influence changes in the dependent variable, but doing so may introduce a correlation between the regressors and the error. In 1991, Arellano and Bond proposed a generalized method of moments approach to address these issues. We use their approach here. The inclusion of two lagged variables was determined so that the residuals are no long autocorrelated on the 2nd order.

Likewise, we modify the market level model to include a lagged dependent variable $\text{Market\_Transaction}_{m,t-1}$

$$\text{Market\_Transaction}_{m,t+1} = \beta_5 + \gamma \text{Market\_Transaction}_{m,t} + \gamma_1 \text{Market\_Transaction}_{m,t-1} +$$

$$\gamma_2 \text{Market\_Transaction}_{m,t-2} + \beta_6 \text{Rating\_Entropy\_Mkt}_{m,t} + \beta_7 \text{Rating\_Entropy\_Mkt}^2_{m,t}$$

$$+ \text{Controls}_{m,t} + \epsilon_{m,t+1}.$$

| | (1) | (2) | (3) |
|---|---|---|---|
| | F.Activity | F.Activity | F.Activity |
| Entropy Others | | 0.743+ | 0.705+ |
| | | (0.057) | (0.070) |
| Entropy(Others)$^2$ | | -1.302* | -1.254* |
| | | (0.039) | (0.041) |
| Entropy(Self) | | | 1.327* |
| | | | (0.004) |
| Entropy(Self)$^2$ | | | -1.695* |
| | | | (0.031) |
| Activity | 0.529* | 0.505* | 0.504* |
| | (0.000) | (0.000) | (0.000) |
| L.Activity | 0.0624* | 0.0377 | 0.0354 |
| | (0.004) | (0.153) | (0.180) |
| Avg | -0.00942 | -0.0701* | -0.0686* |
| | (0.591) | (0.000) | (0.000) |
| Reviews Count | 0.0256* | 0.0179* | 0.0117* |
| | (0.001) | (0.006) | (0.047) |
| Avg(Others) | -0.199+ | -0.200 | -0.167 |
| | (0.077) | (0.111) | (0.167) |
| Experience | -0.0230 | -0.0454 | -0.0548 |
| | (0.680) | (0.415) | (0.309) |
| Price Diff | -0.0329 | 0.115 | 0.114 |
| | (0.742) | (0.291) | (0.299) |
| Market Revenue | 0.00685 | 0.00337 | 0.00376 |
| | (0.240) | (0.618) | (0.581) |
| State Dummies | Yes | Yes | Yes |

**Table 11     Robustness Check: Individual with Dynamic Panel**

## 6.4.   Market Level Alternative Measure of Success

In the analysis of ratings dispersion on local market level performance, we used total quotes accepted by consumers to measure the success of marketplace. We present results using total quotes given out by installers, and it remains consistent, as table 13 shows.

## 6.5.   Text-based Dispersion measure

We used median of cosine distances for measure of dispersion. The mean of cosine distances are consistent, per table 14.

## 6.6.   Excluding Inactive Installers

Although we do not explicitly model the process of installers exiting platform, we are aware of its potential to drive results. We ran a robustness check excluding installers that have been inactive for two month ( making 0 proposals), with results presented in table 15. The first two columns are results excluding these said installers ( cluster standard errors on market level - column (1); individual level - column (2)) . The results are virtually unchanged, especially on the independent variable of interests.

**Authors' names blinded for peer review**

26                    Article submitted to *Manufacturing & Service Operations Management*; manuscript no.

|                          | (1)           | (2)           |
|--------------------------|---------------|---------------|
|                          | F.Transaction | F.Transaction |
| Entropy                  | 1.957***      | 1.050**       |
|                          | (0.000)       | (0.004)       |
| Entropy$^2$              | -2.456***     | -1.609***     |
|                          | (0.000)       | (0.000)       |
| Rating Avg               | -0.207        | -0.141        |
|                          | (0.169)       | (0.476)       |
| Mkt Revenue              | -0.0264       | 0.0403        |
|                          | (0.502)       | (0.051)       |
| Experience(Avg)          | 0.0102        | -0.00118      |
|                          | (0.157)       | (0.882)       |
| Price Diff(Avg)          | 0.0349        | -0.313        |
|                          | (0.865)       | (0.089)       |
| Transaction              |               | 0.0238        |
|                          |               | (0.685)       |
| L.Transaction            |               | -0.00533      |
|                          |               | (0.917)       |
| L2.Transaction           |               | 0.198**       |
|                          |               | (0.001)       |
| Constant                 | 2.015**       |               |
|                          | (0.006)       |               |
| Fixed Effects            | Yes           | No            |
| Weighted State Dummies   | Yes           | Yes           |
| Observations             | 642           | 421           |
| r2                       | 0.726         |               |

**Table 12      Robustness Check Market Level with Dynamic Panels**

## 7.   Discussions

*Average Ratings and Sentiment* : Most the specifications concerning the impact of average ratings captured negative (yet statistically insignificant) effects. Interestingly, the model using sentiment score and Text-based dispersion measures (table 9) have shown more consistent and significant negative coefficients. After we control for other things, being rated higher or viewed more positive is associated with a lower level of activity intensity going forward.

*Noise in Reviews* The individual level analysis pertain to ratings and reviews dispersion all revealed an inverted-U type of impact.

*Noisy Reviews and Marketplace Matching Performance* : from the marketplace's perspective, dispersion in reviews also exhibits an inverted-U type of relationship with market place matches.

*Methodology - text mining* To analyze the reviews texts, we incorporated two text mining methods that 1) - gave reviews texts a one-dimensional sentiment score and 2) utilize word embedding model to measure texts similarity with precision. We demonstrated that the text mining tools are great complement to the quantitative data. To our knowledge, it is the first example of using

| | (1) | (2) | (3) |
|---|---|---|---|
| | F.Activity | F.Activity | F.Activity |
| Entropy | 4.189** | 4.290** | |
| | (0.003) | (0.004) | |
| Entropy # Entropy | -4.066* | -4.147* | |
| | (0.020) | (0.018) | |
| Text-based Entropy | | | 20.57 |
| | | | (0.348) |
| Text-based Entropy # Text-based Entropy | | | -61.58 |
| | | | (0.332) |
| Rating Avg | -0.218 | | |
| | (0.695) | | |
| Sentiment Score | | -0.217 | 0.349 |
| | | (0.651) | (0.601) |
| Mkt Revenue | -0.282** | -0.288** | -0.379** |
| | (0.008) | (0.004) | (0.005) |
| Mkt Revenue # Mkt Revenue | 0.0158* | 0.0161** | 0.0214** |
| | (0.014) | (0.007) | (0.008) |
| Experience(Avg) | 0.0559* | 0.0573* | 0.0740** |
| | (0.017) | (0.013) | (0.002) |
| Price Diff(Avg) | 0.0970 | -0.00130 | -0.000982 |
| | (0.832) | (0.998) | (0.998) |
| Constant | 6.594* | 5.596*** | 3.642 |
| | (0.019) | (0.000) | (0.053) |
| Observations | 746 | 767 | 961 |
| p-values in parentheses | | | |
| ="* p<0.05 | | ** p<0.01 | *** p<0.001" |

**Table 13      Market Level Use Given Quotes( instead of winning quotes)**

deep learning based text-mining models on business settings in the operations management literature. We demonstrate the versatility of deep learning methods as a complement to traditional text-mining methods.

# Appendix. Note and Questions for Our Own References

## A.   Tests for Fixed and Random Effects in (2)

A Individual Level - FE vs RE Hausman Test: Ho: difference in coefficients not systematic

$$chi2(13) = (b - B)'[(V_b - V_B)^( - 1)](b - B)$$

$$= 44.23$$

$$Prob > chi2 = 0.0000$$

Market Level -FE vs. RE Hausman

Test: Ho: difference in coefficients not systematic

$$chi2(18) = (b - B)'[(V_b - V_B)^( - 1)](b - B) = 403.30 \quad Prob > chi2 = 0.0000$$ **PLEASE INSERT A TABLE THAT REPORTS THE TEST RESULTS**

28

**Authors' names blinded for peer review**
Article submitted to *Manufacturing & Service Operations Management*; manuscript no.

|                                   | (1)        | (2)        |
| --------------------------------- | ---------- | ---------- |
|                                   | F.Activity | F.Activity |
| Text-based Entropy Others         | 31.75*     | 4.758      |
|                                   | 0          | -0.675     |
| Text-based Entropy Others $^2$    | -103.5*    | -18.55     |
|                                   | 0          | -0.644     |
| Text-based Entropy (self)         |            | 13.16      |
|                                   |            | -0.284     |
| Text-based Entropy (self)$^2$     |            | -37.23     |
|                                   |            | -0.246     |
| Avg Sentiment Score               | -0.825*    | -1.082+    |
|                                   | -0.027     | -0.05      |
| Avg Sent(Others)                  | -0.199     | -0.221     |
|                                   | -0.445     | -0.404     |
| Reviews Count                     | 0.0484*    | 0.0483*    |
|                                   | 0          | 0          |
| Experience                        | 0.236*     | 0.0946     |
|                                   | -0.002     | -0.282     |
| Price Diff                        | -0.0598    | 0.0299     |
|                                   | -0.38      | -0.884     |
| Market Revenue                    | -0.0163*   | -0.0159+   |
|                                   | -0.012     | -0.073     |
| Constant                          | 0.166      | 1.367      |
|                                   | -0.776     | -0.404     |
| Observations                      | 6210       | 3479       |
| p-values in parentheses           |            |            |
| ="+ p<0.10                        | * p<0.05"  |            |

**Table 14    Robustness Individual with Text-based Dispersion from Mean Distance**

## B.   My Earlier Question

What is the total number of proposals in each year? Which state is number #1 in terms of total wins/total proposals? Which state is worst in terms of total wins/total proposals? Q1) IF THEY OPERATE AT MULTIPLE LOCATIONS, DO THEY PROVIDE THAT INFO ON THEIR PROFILE?

We do not have info if they operated on multiple locations or not. I scrape their headquarter address , with ZIP code info. Q2) WHAT IS THE FORMAT OF THE LOCATION INFO - IS IT A DETAILED ONE WITH A ZIPCODE? PLEASE INCLUDE AN EXAMPLE FOR ME HERE.

Emerald Energy of North Carolina Headquarters 2624 Leighton Ridge Drive, Suite 120 Wake Forest, NC 27587 US

  COULD YOU PLEASE PREPARE THESE TWO GRAPHS? 1) TOTAL NUMBER OF REVIEWS PER INSTALLER - MAX NUMBER OF REVIEW FOR AN IN- STALLER AND HISTOGRAM? 2) NUMBER OF INSTALLERS IN EACH STATE FOR TOP 10 STATES

QUESTIONS: 1) HOW MANY OF THE INSTALLERS DO NOT HAVE A RE- VIEW? 2) WHAT IS THE PERCENTAGE OF THOSE INSTALLERS IN THE LO- CAL MARKET WIN AND SUBMITTED PROPOSALS? 3) WHAT DO YOU AS- SUME ABOUT THEM IN THE EMPIRICAL ANALYSIS?

|  | (1) | (2) | (3) |
|---|---|---|---|
|  | F.Activity | F.Activity | F.Activity |
| Average Rating | -0.551* | -0.524* | -0.489* |
|  | (0.007) | (0.011) | (0.033) |
| Average Ratings(Others) | 0.00733 | 0.000713 | 0.0000629 |
|  | (0.741) | (0.977) | (0.998) |
| Reviews Count | 0.0504* | 0.0479* | 0.0436* |
|  | (0.000) | (0.000) | (0.000) |
| Experience | 0.136* | 0.134* | 0.133* |
|  | (0.010) | (0.012) | (0.012) |
| Price Diff | 0.188* | 0.200* | 0.208* |
|  | (0.036) | (0.027) | (0.021) |
| Market Revenue | -0.00645 | -0.00618 | -0.00553 |
|  | (0.148) | (0.165) | (0.216) |
| Entropy(Others) |  | 0.386+ | 0.358 |
|  |  | (0.085) | (0.108) |
| Entropy(Others)$^2$ |  | -2.243* | -2.312* |
|  |  | (0.000) | (0.000) |
| Entropy(Self) |  |  | 1.428* |
|  |  |  | (0.000) |
| Entropy(Self)$^2$ |  |  | -2.590* |
|  |  |  | (0.000) |
| Constant | 2.788* | 2.868* | 2.995* |
|  | (0.000) | (0.000) | (0.000) |
| Observations | 3472 | 3472 | 3472 |
| r2 | 0.622 | 0.625 | 0.627 |
| aic | 9693.2 | 9670.1 | 9655.4 |
| bic | 9748.6 | 9737.7 | 9735.3 |

**Table 15     Robustness Check Excluding Inactive Installers**

| Variable | VIF | 1/VIF |
|---|---|---|
| Average Rating | 4.59 | 0.22 |
| Reviews Count | 1.24 | 0.81 |
| Average Rating (Others) | 1.40 | 0.71 |
| Entropy(others) | 1.83 | 0.55 |
| Entropy(others)^2 | 1.46 | 0.68 |
| Entropy(self) | 7.10 | 0.14 |
| Entropy(self)^2 | 5.81 | 0.17 |
| Experience | 1.46 | 0.68 |
| Price(difference) | 1.03 | 0.97 |
| Market Condition) | 1.65 | 0.61 |

**Table 16     VIF table: Individual Level**

We didn't need to assume anything. I just computed all the variables the way as we stated.

All installers started with 0 reviews, naturally.

If we look at the observations that are included in the analysis, less than 5 percent of observations have 0 reviews ( mostly due to its newly established).

| Variable | VIF | 1/VIF |
|---|---|---|
| Entropy(Market) | 6.67 | 0.149965 |
| Entropy(Market)$^2$ | 6.21 | 0.160950 |
| Average Rating | 1.58 | 0.632615 |
| Market Revenue | 1.45 | 0.690279 |
| Experience(Avg) | 1.46 | 0.685113 |
| Price Difference | 1.01 | 0.986885 |

**Table 17     VIF Table:Market level**

| Individual | | |
|---|---|---|
| Order | z | prob>z |
| 1 | -9.8283 | 0.00 |
| 2 | -0.66053 | 0.5089 |
| **Market** | | |
| Order | z | prob>z |
| 1 | -2.7882 | 0.01 |
| 2 | 0.04295 | 0.9657 |

**Table 18     Arellano-Bond test for zero autocorrelation in first-differenced errors**

If we look at the end of the panel, only 1 installer has 0 reviews, and have a positive entothers value ( hence is included in the analysis)

**YOU MENTIONED controls are to capture factors that are irrelevant to the rating entropy. Is Experience really irrelevant to the rating entropy??**

# References

Agency IE (2018) Snapshot of global photovoltaic markets; 2018 URL `http://www.iea-pvps.org/fileadmin/dam/public/report/statistics/IEA-PVPS_-_A_Snapshot_of_Global_PV_-_1992-2017.pdf`.

Boris (2018) How we determine valuations for marketplaces. URL `https://versionone.vc/how-we-determine-valuations-for-marketplaces/`.

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.

Davidson M (2019) What neighborhood is this? estimating material based neighborhood boundaries using single-family homes. *Estimating Material Based Neighborhood Boundaries Using Single-Family Homes (November 13, 2019)* .

Dellarocas C, Zhang XM, Awad NF (2007) Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive marketing* 21(4):23–45.

Devlin J, Chang MW, Lee K, Toutanova K (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* .

Dorsey J (2019) Access to alternatives: Increasing rooftop solar adoption with online platforms. *Kelley School of Business Research Paper* .

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

Galston E (2017) Anatomy of a managed marketplace. URL `https://techcrunch.com/2017/05/25/anatomy-of-a-managed-marketplace/`.

Haddad M, Kleiner B (2015) Consumer goods industry: Challenges within the online marketplace. *Ethics & Critical Thinking Journal* 2015(1).

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

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.*

INC E (????) New report finds solar tariffs not helping american-made solar panels compete, and costs of solar continuing to fall. URL `https://www.energysage.com/press/energysage-marketplace-intel-report-8`.

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).

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, ISSN 1540-6261, URL `http://dx.doi.org/10.1111/j.1540-6261.2010.01625.x`.

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

Marinovic I (2015) The credibility of performance feedback in tournaments. *Journal of Economics & Management Strategy* 24(1):165–188.

Pyper J (2018) California's rooftop solar mandate wins final approval. URL `https://www.greentechmedia.com/articles/read/california-solar-roof-mandate-wins-final-approval`.

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

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.

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.

US Census Bureau (2019) Gazetteer files. URL `https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.2018.html`.

Wang F, Liu X, Fang EE (2015) User reviews variance, critic reviews variance, and product sales: An exploration of customer breadth and depth effects. *Journal of Retailing* 91(3):372–389.

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

Zimmermann S, Herrmann P, Kundisch D, Nault BR (2018) Decomposing the variance of consumer ratings and the impact on price and demand. *Information Systems Research* 29(4):984–1002.