Submitted to *Manufacturing & Service Operations Management*
manuscript

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.

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 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

—-[OLD VERSION]

Solar cells, also called photovoltaic cells, convert sunlight directly into electricity without carbon emissions. Today, electricity from solar cells has become competitive in many regions and photovoltaic systems are being deployed at large scales to help power the electric grid (NREL).

Solar energy is blooming in the US and the world. It is one of the fastest growing energy generating technology with a dazzling 34% growth worldwide in 2017 (Agency 2018). Just 6% of American household have already installed solar panels at home with another 46% say they have given serious thought to adding solar panels at their home in the past year (CITE kennedy thigpen

2019).Solar PV capacity increased by an annual rate of 50% in decade and residential solar is forecasted to grow 25% per year (Weaver 2019, SEIA); with an even larger upside in the U.S after the passing of California Solar mandate (Pyper 2018).

Online marketplaces is an innovative business model that has shown to ease the rooftop solar panel adoption process. It serves as an intermediary which connects buyers and made the whole process more transparent (Dorsey 2019). There is an increasing trend of installing rooftop panels through online market places. Consumer interest doubled in 11 states between 2017 to 2018, according to an analysis of website traffic (INC).

In building an online marketplace, online reviews is considered an essential functionality. Studies have shown that reviews have significant impact on customers' decision making process, especially for products and services that entail searching and experiencing attributes (Zimmermann et al. 2018).

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)).

In this paper, different from this literature, our primary goal is to study the impact of dispersion of ratings 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.

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 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).

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

1.1. 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.

1.2. 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.

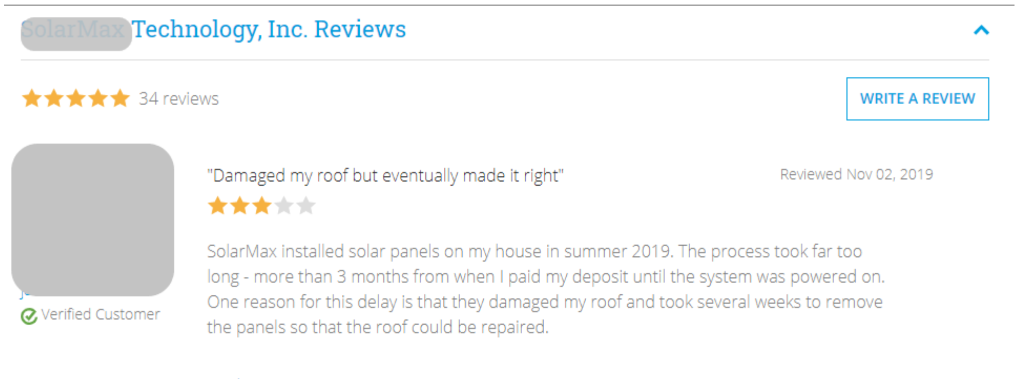

Figure 1 **A sample customer review on the marketplace.**

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.

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

	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

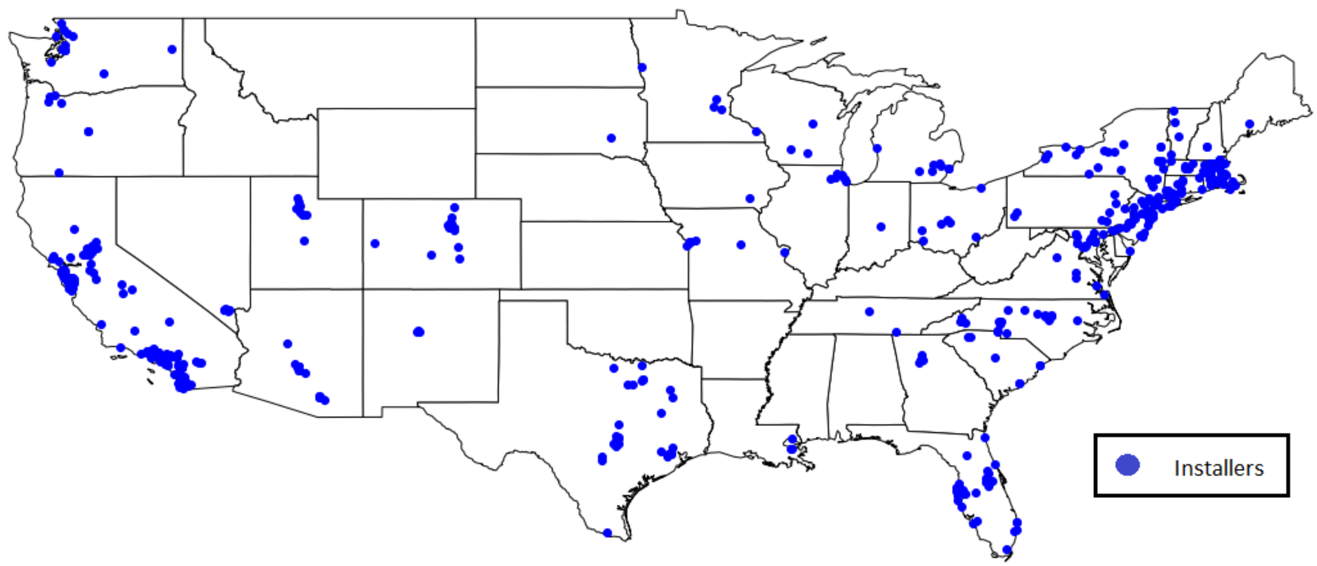

Figure 2 Installers in our data set

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 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).

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. **Is the previous sentence true?** 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.

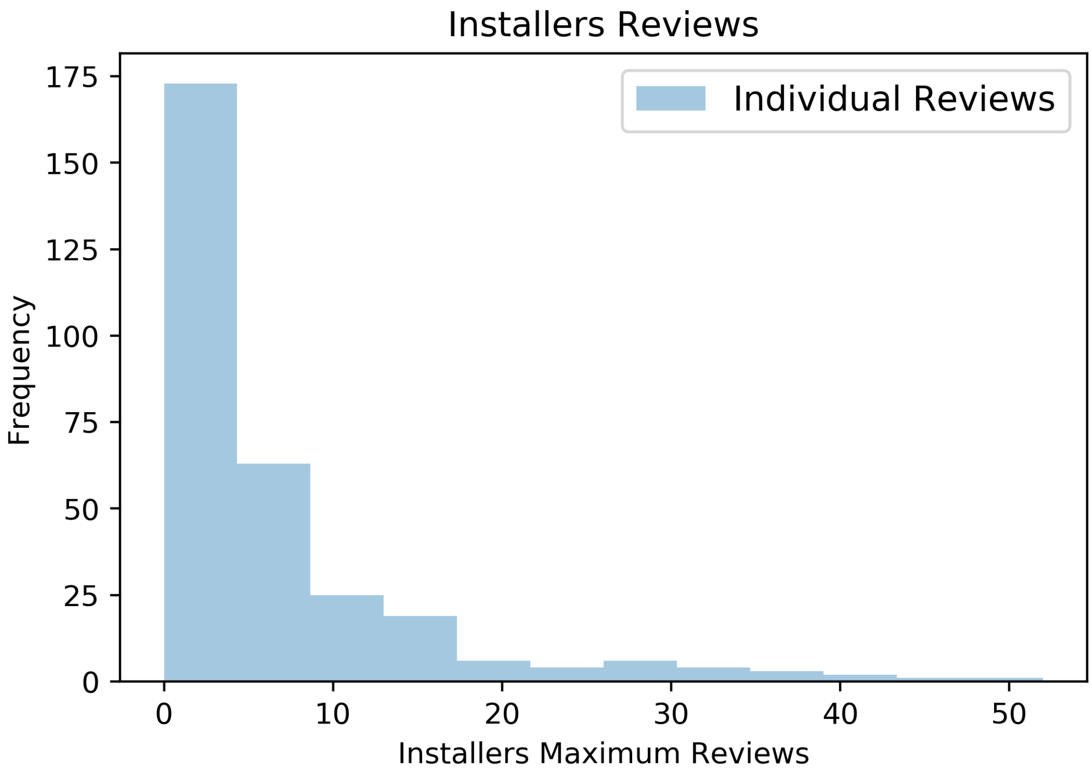

Figure 3 **Reviews Histogram by Installer**

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**

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

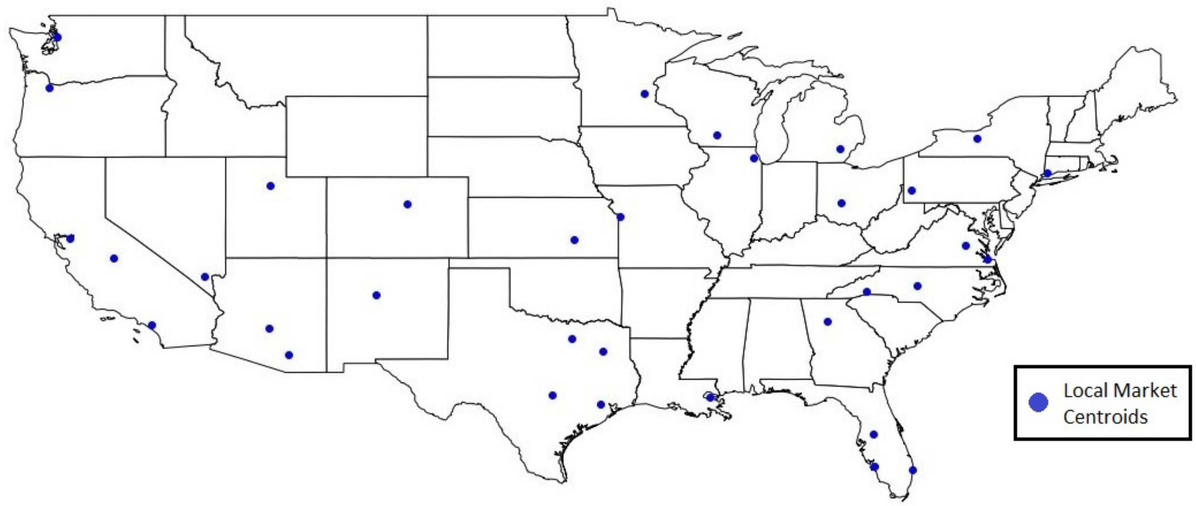

Figure 4 **Local Market Centroids**

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 $\text{Entropy\_Self}_{i,t}$ 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 (**and including?**) month t. Recalling that market boundaries are defined as in Section 2.2, the second variable $\text{Entropy\_Others}_{i,t}$ 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 $\text{Entropy\_Mkt}_{m,t}$ 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.

Explain why we did not use standard deviation /variance to measure dispersion.

Please include the correlation between variance and average of reviews here.

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 word embedding 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 has adopted this method in October 2019 (REFERENCE). To the best of our knowledge, there is 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 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 Empirical 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 $\text{Installer\_Activity}_{i,m,t}$ 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. Ratings-based Model & Controls

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

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

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

The regression (2) includes various additional 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 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 $\text{Price\_Difference}_{i,t}$ 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 $\text{Price\_Difference}_{i,t}$ 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 **IS THIS CALCULATED FOR EACH MONTH?**. We control for the self average rating of each installer i $\text{Avg\_Rating\_Self}_{i,t}$ as well as the average ratings of its competitors $\text{Avg\_Rating\_Others}_{i,t}$ in the market for month t. We also control for the experience of the installer by considering the variable $\text{Experience}_{i,t}$ 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 $\text{Market\_Revenue}_{m,t}$ 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 $Review\_Counts_{i,t}$.

PLEASE VERIFY THE DEFINITION OF Total_Installations.

	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

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

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

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. **PLEASE UPDATE VARIABLE NAMES AND TABLE 3**

3.2. Ratings-based Results

Table **??** presents results estimated by three panel regression models based on (2). Columns (1) through (3) of Table **??** 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 are estimated by considering only some of these variables in the regression.

Our estimates in Table **??** 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 β_1 in (2) is found to be positive and significant in column(3) of Table **??**), and the second-order effect (β_2 in (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.

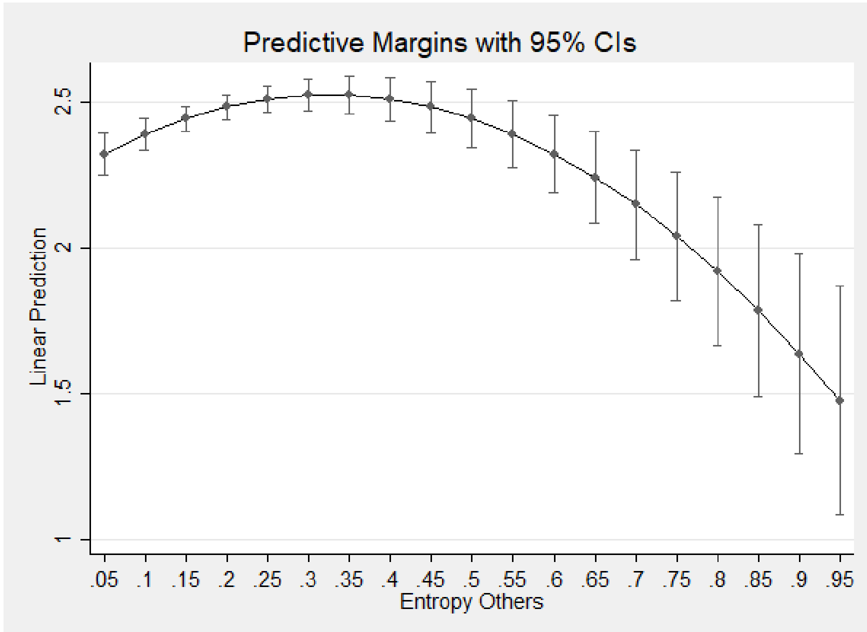

Figure 5 Marginal Impact of the Entropy of Other Installers' Ratings (in the same market) on the Installer's Activity Level

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 from the column (3) of Table **??**. As is apparent from these figures, the installer's activity level first increases then decreases with the rating dispersion, and that is true both when own ratings and its com, regardless of them being based on own ratings or its competitors' ratings.

4. Local Markets

Table

We now move to discuss the ratings dispersion on total transactions on local market level. The results are presented in table 6. Column (1) and column (2) used fixed effect and random effect, respectively.

Signal :The results on the local market level also suggest that after controlling for local market conditions, installer experience, price, the 'signal' portion of the ratings are not significantly associated with the market level performance.

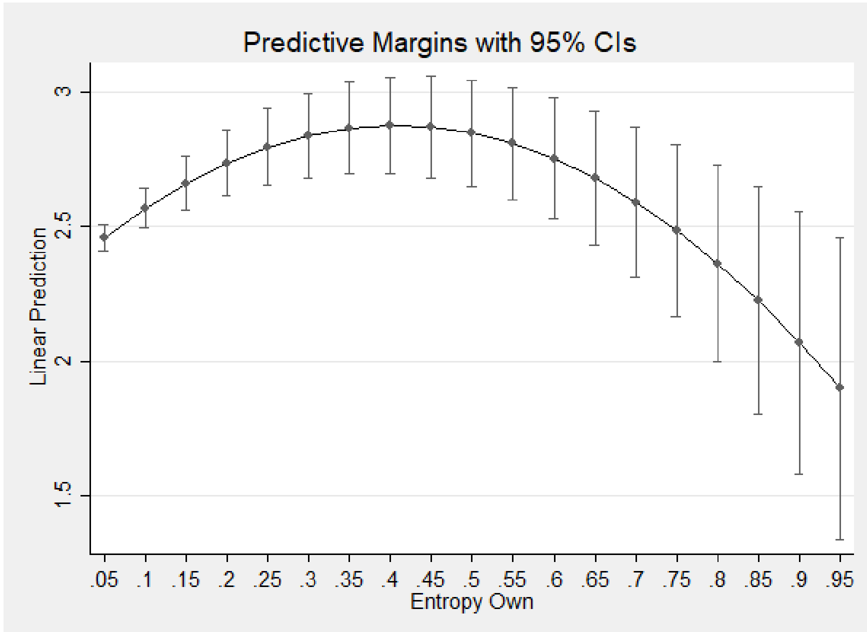

Figure 6 **Marginal Impact of the Entropy of the Installer's Own Ratings on its Activity Level**

	(1)	(2)	(3)	(4)
	F.Transaction	F.Transaction	F.Activity	F.Activity
Rating Avg	10.00**	14.90*	21.31**	22.65**
	(0.005)	(0.011)	(0.004)	(0.001)
Rating Avg ^{2}	-1.140**	-1.694**	-2.458**	-2.613**
	(0.004)	(0.009)	(0.004)	(0.001)
Mkt Revenue	-0.163*	-0.135**	-0.381**	-0.364**
	(0.018)	(0.005)	(0.002)	(0.001)
Experience(Avg)	0.0267*	-0.0203	0.0820**	0.0768***
	(0.029)	(0.499)	(0.001)	(0.000)
Price Diff(Avg)	0.270	0.0177	0.107	-0.0456
	(0.074)	(0.887)	(0.752)	(0.873)
Constant	-19.67*	-31.66*	-39.74*	-44.98**
	(0.011)	(0.015)	(0.012)	(0.003)
Observations	990	990	990	990
p-values in parentheses				
=”* p<0.05	** p<0.01	*** p<0.001”		

Table 5 **Market Level with Only Ratings Average**

Noise : The '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 proposals accepted, as reflected in the coefficient estimates being positive and statistically significant. We also note that the second order effect is negative as the coefficient estimates associated with the square term is negative and statistically significant. We further illustrate this point with a margins plot using coefficients generated from estimates in column X in figure 8.

	(1)	(2)
	F.Transaction	F.Transaction
Entropy	1.892**	3.772***
	(0.004)	(0.000)
Entropy^2	-1.893*	-4.057***
	(0.011)	(0.000)
Rating Avg	-0.0518	-0.0438
	(0.824)	(0.853)
Mkt Revenue	-0.119	-0.0815*
	(0.053)	(0.031)
Experience(Avg)	0.0179	-0.0359
	(0.074)	(0.172)
Price Diff(Avg)	0.313	0.0482
	(0.125)	(0.787)
Constant	1.902	0.264
	(0.105)	(0.810)
Observations	746	746
p-values in parentheses		
=”* p<0.05	** p<0.01	*** p<0.001”

Table 6 **Market Level Use Winning Quotes and Quantitative Dispersion Measures**

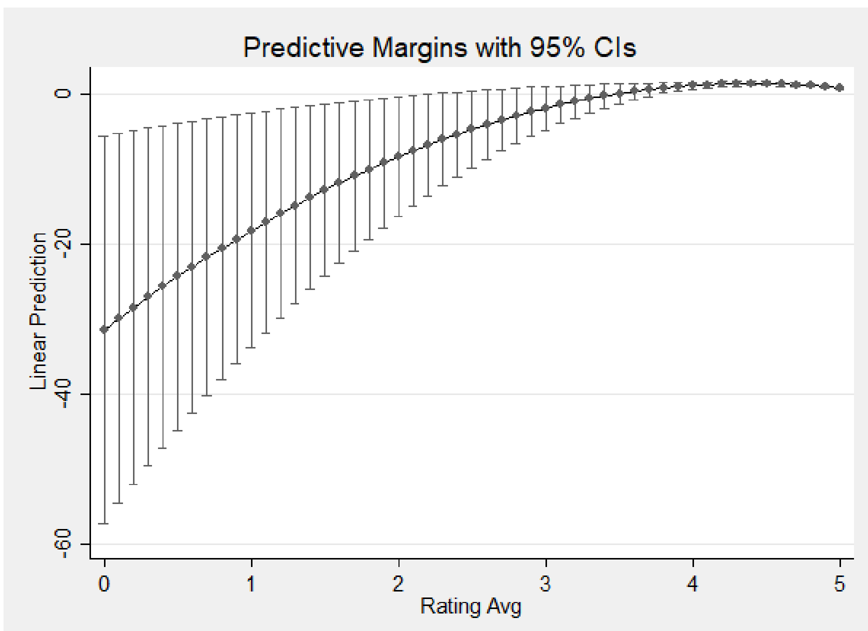

Figure 7 **Marginal Impact of Average Ratings on Market Level Matching**

The relationship we found in this section is similar to what we presented earlier. The important distinction is that we are looking at local market at an entire performance unit and measuring the dispersion of ratings on the market level.

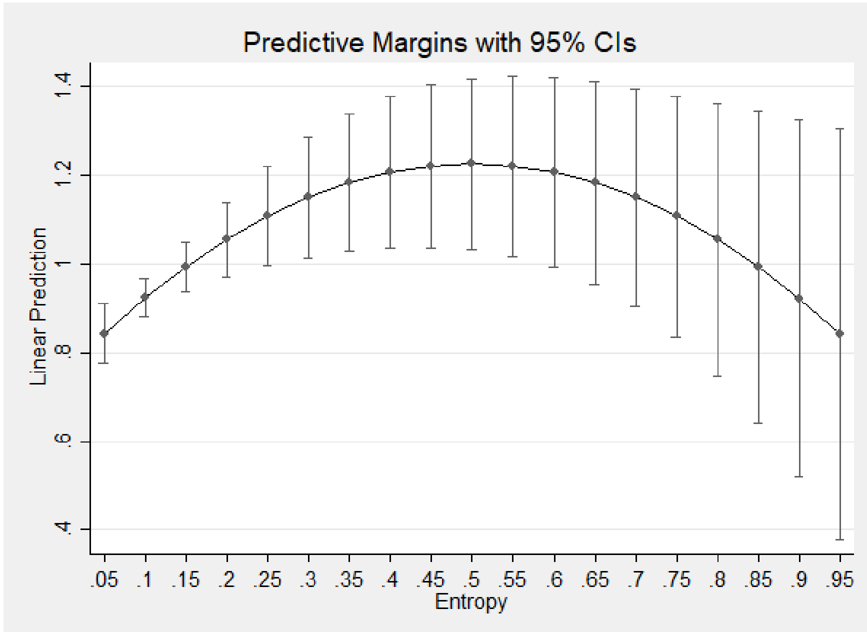

Figure 8 Marginal Impact of Market Reviews Entropy of Reviews on Market Level Matching

5. Text Mining

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

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, is A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text (Hutto and Gilbert 2014). 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: The above review received a compound sentiment score of 0.8622 and a five-star rating.

Example 2: This received -0.7184 and an one-star rating.

We apply this method to assign a sentiment score for every piece of reviews texts. Overall, the sentiment score correlates with ratings signficantly($corr = 0.8239$). It is further illustrated in

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Text-based Entropy Others	25.92*	25.92*	-4.913	-4.913
	(0.003)	(0.001)	(0.663)	(0.626)
Text-based Entropy Others ^{2}	-86.97*	-86.97*	0.206	0.206
	(0.001)	(0.001)	(0.995)	(0.995)
Text-based Entropy(self)			11.25	11.25
			(0.231)	(0.163)
Text-based Entropy(self) ^{2}			-30.79	-30.79
			(0.167)	(0.113)
Avg	-0.629	-0.629	-1.299*	-1.299*
	(0.382)	(0.330)	(0.001)	(0.008)
Avg(Others)	-0.0879	-0.0879	-0.103	-0.103
	(0.633)	(0.559)	(0.576)	(0.512)
Avg ^{2}	0.0670	0.0670	0.167*	0.167*
	(0.554)	(0.495)	(0.011)	(0.025)
Reviews Count	0.0482*	0.0482*	0.0456*	0.0456*
	(0.001)	(0.000)	(0.001)	(0.000)
Experience	0.292*	0.292*	0.110	0.110
	(0.018)	(0.002)	(0.437)	(0.285)
Price Diff	-0.0468	-0.0468	0.0295	0.0295
	(0.540)	(0.628)	(0.880)	(0.861)
Market Revenue	-0.0152*	-0.0152*	-0.0164+	-0.0164*
	(0.025)	(0.037)	(0.084)	(0.027)
Constant	1.834	1.834	4.959*	4.959*
	(0.304)	(0.194)	(0.032)	(0.001)
Observations	5602	5602	3112	3112
p-values in parentheses				
=”+ p<0.10	* p<0.05”			

Table 7 Individual Level Use Average Rating and Text-based dispersion

figure 9 where the scatterplot of all reviews' sentiment scores and ratings is presented. Follow the same process as discussed in earlier section, we construct variables $AvgSent_{i,t}$ in place of $Avg_{i,t}$, $AvgSentOthers_{i,m,t}$ in place of $AvgOthers_{i,m,t}$ to represent individual ratings in individual level analysis, $AvgSentMkt_{m,t}$ in place of $AvgMkt_{m,t}$ in market level analysis.

5.2. Measures of ratings dispersion

We constructed variables that measures variations in reviews to complement dispersions in rating. We show that these two measures are positively correlated, exhibits similar connection with installer and market level activity intensity, and also complement each other.

5.2.1. Capture Dispersion in Texts with language model BERT In addition to ratings, We want to leverage the rich information in the reviews texts. We hypothesize that the *dispersion* in reviews texts shall also exhibit similar effect as the ratings as well as positively correlated with the entropy. A set of all 5 star reviews with praises might contain less information

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Entropy Others	1.654*	1.654*	1.638*	1.638*
	(0.016)	(0.015)	(0.018)	(0.012)
Entropy Others ^{2}	-2.532*	-2.532*	-2.575*	-2.575*
	(0.002)	(0.009)	(0.002)	(0.006)
Entropy Own			2.711*	2.711*
			(0.002)	(0.005)
Entropy Own^2			-3.367*	-3.367*
			(0.017)	(0.016)
Avg Sentiment Score	-1.271*	-1.271*	-1.283*	-1.283*
	(0.004)	(0.011)	(0.007)	(0.012)
Avg Sentiment Score ^{2}	0.632	0.632	0.853+	0.853+
	(0.283)	(0.236)	(0.082)	(0.080)
Avg Sent(Others)	0.124	0.124	0.111	0.111
	(0.627)	(0.594)	(0.657)	(0.628)
Reviews Count	0.0538*	0.0538*	0.0474*	0.0474*
	(0.000)	(0.000)	(0.000)	(0.000)
Experience	0.213*	0.213*	0.201*	0.201*
	(0.012)	(0.008)	(0.016)	(0.010)
Price Diff	0.0816	0.0816	0.0819	0.0819
	(0.555)	(0.557)	(0.521)	(0.542)
Market Revenue	-0.0174*	-0.0174*	-0.0162+	-0.0162*
	(0.050)	(0.015)	(0.083)	(0.025)
Constant	2.390*	2.390*	2.277*	2.277*
Observations	4562	4562	4562	4562
p-values in parentheses				
=”+ p<0.10	* p<0.05”			

Table 8 Individual Analysis Use Sentiment Score and Entropy

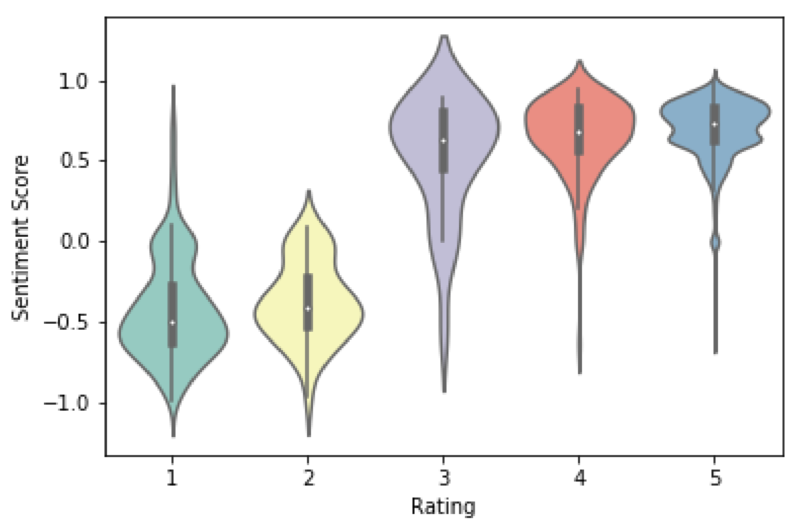

Figure 9 Sentiment Scores and Ratings

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Text-based Entropy Others	25.40*	25.40*	-10.30	-10.30
	(0.001)	(0.001)	(0.316)	(0.311)
Text-based Entropy Others^2	-84.20*	-84.20*	24.66	24.66
	(0.000)	(0.001)	(0.407)	(0.447)
Text-based Entropy(self)			11.27	11.27
			(0.263)	(0.179)
Text-based Entropy(self) ^{2}			-28.49	-28.49
			(0.235)	(0.157)
Avg Sentiment Score	-0.797*	-0.797	-1.100+	-1.100*
	(0.031)	(0.101)	(0.058)	(0.045)
Avg Sentiment Score^2	0.263	0.263	1.552*	1.552*
	(0.621)	(0.638)	(0.012)	(0.015)
Avg Sent(Others)	-0.147	-0.147	-0.204	-0.204
	(0.599)	(0.495)	(0.482)	(0.476)
Reviews Count	0.0482*	0.0482*	0.0492*	0.0492*
	(0.001)	(0.000)	(0.000)	(0.000)
Experience	0.267*	0.267*	0.0924	0.0924
	(0.018)	(0.002)	(0.463)	(0.320)
Price Diff	-0.0641	-0.0641	0.0343	0.0343
	(0.350)	(0.481)	(0.861)	(0.837)
Market Revenue	-0.0156*	-0.0156*	-0.0170+	-0.0170*
	(0.016)	(0.031)	(0.072)	(0.021)
Constant	0.530	0.530	2.541	2.541*
	(0.403)	(0.450)	(0.113)	(0.034)
Observations	5722	5722	3179	3179
p-values in parentheses				
=”+ p<0.10	* p<0.05”			

Table 9 Use Sentiment score and Text-based Entropy

than a mix of 1, 2 and 5 stars. This is reflected in entropy as the later will have a higher entropy. We aim to design a measure that captures a similar concepts on texts. To achieve that goal of measuring reviews dispersions in texts, we combine the methods inspired by Hoberg and Phillips (2016), tweak it to apply to our data structure, and updated it with a word embedding model called BERT , which we will describe later.

Hoberg and Phillips' work involves measuring the similarity between two pieces of texts. In their case, they measure the distance of the two pieces of business descriptions from 10-k form and take $1 - distance$ to represent similarities between two business entities. Their methods include: 1) Vectorize each piece of text based on the distinct words it contains. 2) Normalize the vectors to unit length. and 3) Use the Cosine similarity to measure how similar are two word vectors. It is called cosine similarity because it measures the angle between the two vectors that represents the texts. If the angle is 0, their similarity shall be 1 and distance be 0. The cosine similarity between

	(1)	(2)	(3)	(4)	(5)	(6)
	F.Tran	F.Tran	F.Tran	F.Tran	F.Tran	F.Tran
Entropy	1.916**	3.763***				
	-0.002	0				
Entropy^2	-1.892**	-4.002***				
	-0.008	0				
Text-based Entropy			10.57	17.88	9.512	19.92
			-0.12	-0.132	-0.258	-0.096
Text-based Entropy^2			-27.66	-56.75	-31.38	-62.58
			-0.123	-0.122	-0.21	-0.095
Rating Avg			-0.148	-0.248		
			-0.447	-0.315		
Sentiment Score	-0.242	0.00699			-0.102	0.277
	-0.245	-0.973			-0.704	-0.257
Mkt Revenue	-0.119*	-0.0787*	-0.127*	-0.0908*	-0.147*	-0.102*
	-0.04	-0.039	-0.017	-0.035	-0.013	-0.036
Experience(Avg)	0.0173	-0.0355	0.0197*	-0.0263	0.0230*	-0.0235
	-0.087	-0.176	-0.035	-0.35	-0.039	-0.391
Price Diff(Avg)	0.352	0.118	-0.0412	-0.229	0.108	-0.177
	-0.093	-0.5	-0.866	-0.28	-0.639	-0.369
Constant	1.789***	0.0537	1.869	0.32	1.218	-1.218
	0	-0.702	-0.077	-0.845	-0.095	-0.151
Observations	767	767	928	928	961	961
p-values in parentheses						
=”* p<0.05	** p<0.01	*** p<0.001”				

Table 10 Market Level Use Use Sentiment Score and Text-based Dispersion

the two vectors is calculated as follows:

- Cosine *Similarity* between V_1 and $V2 = (V_1 \cdot V_2)$
- Cosine *Distance* between V_1 and $V2 = 1 - (V_1 \cdot V_2)$

We incorporate the aforementioned cosine distance concept to measure dispersion in sets of reviews texts. It is achieved by enumerating all pairwise distances of reviews and take its statistical median. For example, on a set of 10 reviews texts pieces, we have 45 (45=$\binom{10}{2}$) pair-wise distances. We then compute the median distances of these 45 similarity scores, denoted as TD to represent the **Text Dispersion**. If the 10 pieces of texts are dissimilar from each other, they contain richer information and the median of these 45 distances data shall be higher; and vice versa.

Similar to ratings entropy, we compute text-based dispersion on 3 different scopes and use them as Independent Variables of interest:

- 1. Text-based Dispersion for one's own reviews up to month t is computed on the N_{it} reviews available up to month t. It is calculated by computing the $N_{it} \times (N_{it} - 1)/2$ cosine distance pairs and take the 50 percentile, which is denote as $TD_{self,i,t}$ (TD: Text-based Dispersion)

- 2. Text-based Dispersion for others' review up to month t is computed on the $N_{i,others,t}$ reviews available up to month t that is in focal installer i's local market. It is calculated by computing the $N_{i,others,t} \times (N_{i,others,t} - 1)/2$ cosine distance pairs and take the 50 percentile, which is denote as $TD_{Others,i,t}$

We also compute the text-based dispersion for every *market-month*:

- 3. Text-based Dispersion for a market m at month t is computed on the $N_{m,t}$ reviews available up to month t. Take the $N_{mt} \times (N_{mt} - 1)/2$ cosine distance pairs and take the 50 percentile and denote it as $TD_{market,i,t}$

We now describe the process we took to *vectorize* the review texts. In our study, we used a BERT word embedding model (Devlin et al. 2018). BERT is short for Bidirectioanl Encoder Representations from Transformers (BERT). It is a natural language processing model that transforms texts into numeric vectors while also preserve the semantic meaning of the texts. It is getting widely applied in research and industry application such as Google Search. It belongs to the category of NLP methods called word embedding. We perform word embedding on the texts before computing distance.
Some earlier literature such as Hoberg and Phillips (2016) used a 'Bag of Words' approach. The 'Bag of Words' approach create one-hot vectors to encodes the appreances of words(or combined counter vector with a tf-idf (term-frequency-inverse document frequency) weighting scheme as used by Loughran and McDonald (2011)). It was an appropriate application for formal financial documents such as 10-K forms. In our application, we are dealing with texts that are informal writings and often with emotions expressed in the text. Simply capturing word frequencies will not be enough if similar emotions can be expressed with synonymous words. We want to produce vectors that will preserve the information and sentiment of the reviews texts despite use of synonyms and/or different styles. We use the following exampls to demonstrate the power of an advanced language model.
1. Understanding the semantics.
For example, consider 3 sentences:
Sentence 1: they did a good job.
Sentence 2: they did an awful job.
Sentence 3: they did a great job.

We want the distance between sentence 1 and 3 to be closer than the distance between 2 and 3 or 1 and 2. Word embedding method enables just that. Word embedding will project "good" and "great" to vectors that are closer together. Without word embedding, the distance between the 3 sentences will be similar (with tf-idf weighting) or the same (without tf-idf weighting, simply use a counter vectorizer).

With the BERT model vectorization,

Similarity between sentence 1 and 2: 0.9134093016230975

Similarity between sentence 2 and 3: 0.9053232267859165

Similarity between sentence 1 and 3: 0.9737446020998256

$S(1,3) > S(1,2) > S(2,3)$

2. Take word ordering into account.

Another advantage that language model has over bag-of-words approach is its consideration of word ordering in sentence meaning. For example, we know that the two sentences ' The food was good, not bad at all' and 'The food was bad, not good at all' have the exact opposite meaning. A 'bag of words' vecotrization will not capture this, but a language model with directional encoding will. For example, consider the 3 sentences:

sentence1= The food was good, not bad at all.

sentence2=The food was great, not bad at all.

sentence3=The food was bad, not good at all.

Similarity between sentence 1 and 2 : 0.978986918926239

Similarity between sentence 2 and 3 : 0.9468497633934021

Similarity between sentence 1 and 3 : 0.9698476195335388

$S(1,2) > S(1,3) > S(2,3)$, the similarity between sentence 2 and 3 is the lowest.

We used the python library via spaCy v2.1 to implement BERT. We converted every piece of reviews text, regardless of its original length, into a numeric vector of shape 768 $\times$ 1, performed calculation on pairwise cosine distances and derived statistical means for every installer-month or market-market as previous mented. The end result is a set of variables representing the dispersion in texts, denoted as $TD_{self,i,t}, TD_{Others,i,t}, TD_{market,m,t}$ that are parallel to the Entropy measures $ENT_{self,i,t}, ENT_{Others,i,t}, ENT_{m,t}$

5.3. Analysis using Variables Derived From Text Mining

We replaced measures derived from quantitative ratings (average ratings) with the average sentiment scores and replace ratings entropy with text-based dispersion and re-run both individual and market level analysis. The results are presented in table 8 and table 9 . Likewise we run the same regression on the market level data and presented the results in table 10. We observe the same type of inverse U shape for the marginal impact of text-based dispersion. This result comes at no surprise as the two measures of ratings dispersion are correlated significantly, although the magnitude of correlation isn't very high (). We found that the even after we include both

6. Market-Level Analysis

We next perform the analysis on a market level. We analyze the connection between market level ratings dispersion and the market level outcomes. We use a regression model with fixed effects and clustered standard errors on the local market level.

To measure the success of the market, we use the total number of accepted quotes. There are several reasons that we use accepted quotes as the performance metrics: 1. The goal of the market place is to help customers connect with installers. 2. The market itself, just like many other market place, is also evaluated by the transaction volume in a business sense.

We create the dependent variable of interests using the following data transformation. For every local market m, we sum up the total number of quotes accepted per that month ($QuotesWon_{imt}$) for every installers i on that market, and take the log transform.

$$SumQuotes_{m,t} = \sum_{i \in m} QuotesWon_{i,m,t}$$
$$MarketActivity_{m,t} = \log(SumQuotes_{m,t} + 1)$$

By doing that, we convert the installer-monthly level panel data from previous section to a market-monthly level panel data so that we can exploit the variations on market level reviews dispersion to identify their impact on local market outcomes.

Using the indexes m for local market, t for month, the following regression equation is used to estimate the impact of ratings dispersion on the local market on the local market performance metrics.

$$MarketActivity_{m,t+1} = \beta Ent_{m,t} + \beta Ent_{m,t}^2 + Controls + \epsilon_{mt} \tag{3}$$

Where $MarketActivity_{m,t+1}$ indicate the log of the total number of proposals accepted on market m in month $t+1$, and the model link it to the $Ent_{i,m,t}$ - Entropy of reviews from all installers on that local market.

6.0.1. Control Variables

We use the following control variables for the market level analysis

State. There are 33 different state represented in the data set, so we created 33 state dummies. Some market span across more than one state. In that case, we weighted state dummy with the percentage.

Experience: similar to individual level experience. We created $AvgExp_{m,t}$ variable to represent the average experience of installers on the local market.

Average Ratings: similar to individual level analysis, we use the $AvgRating_{m,t}$ to control for the average rating of installers on the local market. A higher $AvgRating$ may improve give the local market a boost across the board. **Market condition**: Similar to individual analysis, we use the total monthly revenues from that market to control for market conditions. **Price**: Follow the individual analysis, we look at the difference of average unit price between marketplace and off-marketplace. We use $PriceDiff_{m,t}$ to denote this variable.

$$\log ZipRev_{m,t} = \log \sum_{j \in m} Rev_{j,t}$$

The total number of reviews on the market We use

$$SumReviews_{m,t} = \sum_{i \in m} Reviews_{i,m,t}$$

7. Robustness Check

7.1. 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

$$ActInt_{i,m,t+1} = \delta pro_{i,t} + \beta_1 Ent_{i,m,others,t} + \beta_2 Ent^2_{i,m,others,t} + controls + \epsilon_{i,m,t} \tag{4}$$

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 β (CITE ECONOMETRIC stuff).

7.2. 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.

7.3. Dynamic Panel model

In our main analysis we include both 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 variable (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:

$$ActInt_{i,m,t+1} = \gamma ActInt_{i,m,t-1} + Ent_{i,m,others,t} + Ent^2_{i,m,others,t} + controls + \epsilon_{i,m,t} \tag{5}$$

$$ActInt_{im,t+1} = \gamma Ent_{im,t-1} + \beta_3 + \beta_4 Ent_{i,self,t} + \beta_5 Ent^2_{i,self,t} + Controls + \epsilon_{imt} \tag{6}$$

$$ActInt_{im,t+1} = \gamma Ent_{im,t-1} + \beta_6 + \beta_7 Ent_{i,self,t} + \beta_8 Ent^2_{i,self,t} + \beta_9 Ent_{im,others} + \beta_{10} Ent^2_{im,self,t} + Controls + \epsilon_{imt} \tag{7}$$

We expect γ estimates to be positive. The results are still consistent as the β coefficients associated with $Ent_{others}(Ent^2_{others})$ are still positive (negative) as presented in table 11 and 12.

Likewise, we modify the market level model to include a lagged dependent variable $MarketActivity_{m,t-1})$

$$MarketActivity_{m,t+1} = \gamma MarketActivity_{m,t-1} + \beta Ent_{m,t} + \beta Ent^2_{m,t} + Controls + \epsilon_{mt} \tag{8}$$

and the results, presented in table 13, are still consistent.

7.4. 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 14 shows.

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Avg	-0.655	-0.655	-0.508	-0.508
	(0.345)	(0.283)	(0.204)	(0.211)
Avg ^{2}	0.0471	0.0471	0.0375	0.0375
	(0.660)	(0.614)	(0.572)	(0.563)
Reviews Count	0.0480*	0.0480*	0.0273*	0.0273*
	(0.000)	(0.000)	(0.000)	(0.001)
Avg(Others)	-0.0285	-0.0285	-0.0491	-0.0491
	(0.873)	(0.876)	(0.754)	(0.717)
Entropy Others	1.762*	1.762*	1.485*	1.485*
	(0.020)	(0.012)	(0.009)	(0.006)
Experience	0.212*	0.212*	0.141+	0.141+
	(0.021)	(0.018)	(0.083)	(0.056)
Price Diff	0.0861	0.0861	0.0120	0.0120
	(0.569)	(0.557)	(0.908)	(0.913)
Market Revenue	-0.0169+	-0.0169*	-0.0136+	-0.0136*
	(0.074)	(0.023)	(0.100)	(0.024)
Entropy Others ^{2}	-2.626*	-2.626*	-2.148*	-2.148*
	(0.001)	(0.005)	(0.000)	(0.003)
L.Activity			0.324*	0.324*
			(0.000)	(0.000)
Constant	4.062*	4.062*	3.128*	3.128*
	(0.004)	(0.004)	(0.004)	(0.003)
Observations	4190	4190	4047	4047
p-values in parentheses				
=”+ p<0.10	* p<0.05”			

Table 11 Robustness Check Add Lagged DV

7.5. Text-based Dispersion measure

We used median of cosine distances for measure of dispersion. The mean of cosine distances are consistent, per table 15.

7.6. 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 16. 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.

8. Discussions

The impact of signal and noise in performance reviews on action-takers are perplexing. Prior literature ...signal... , the collective role of signal and noise have not been empirically investigated. Our study makes several important contributions:

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Avg	-1.170+	-0.862+	-1.155+	-0.850+
	(0.068)	(0.054)	(0.060)	(0.055)
Avg ^{2}	0.130	0.0944	0.129	0.0926
	(0.180)	(0.203)	(0.170)	(0.207)
Reviews Count	0.0452*	0.0254*	0.0423*	0.0230*
	(0.000)	(0.001)	(0.000)	(0.002)
Avg(Others)	-0.0310	-0.0611	-0.0273	-0.0435
	(0.857)	(0.672)	(0.887)	(0.795)
Experience	0.211*	0.141+	0.207*	0.139+
	(0.019)	(0.077)	(0.023)	(0.086)
Price Diff	0.0863	0.0160	0.0928	0.0188
	(0.554)	(0.875)	(0.516)	(0.849)
Market Revenue	-0.0163	-0.0131	-0.0160	-0.0130
	(0.107)	(0.135)	(0.109)	(0.134)
Entropy Own	2.616*	2.009*	2.676*	2.061*
	(0.006)	(0.008)	(0.004)	(0.004)
Entropy Own ^{2}	-3.090*	-2.425*	-3.301*	-2.602*
	(0.042)	(0.024)	(0.020)	(0.007)
L.Activity		0.322*		0.320*
		(0.000)		(0.000)
Entropy Others			1.740*	1.478*
			(0.022)	(0.009)
Entropy Others^2			-2.674*	-2.204*
			(0.001)	(0.000)
Constant	4.698*	3.649*	4.519*	3.437*
	(0.001)	(0.001)	(0.002)	(0.003)
Observations	4190	4047	4190	4047
p-values in parentheses				
=”+ p<0.10	* p<0.05”			

Table 12 Robustness Check with Lagged Variable and Own Entropy

Average (Signal) : 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 8 and 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.

EntOthers (Noise) The individual level analysis pertain to *EntOthers* covariates all revealed an inverse-U shape impact.

EntSelf (Noise)

Market Level Impact

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

	(1)	(2)	(3)	(4)
	F.Tran	F.Tran	F.Tran	F.Tran
Entropy	2.106*	2.106*	1.682*	1.682*
	(0.000)	(0.003)	(0.000)	(0.001)
Entropy ^{2}	-2.163*	-2.163*	-1.893*	-1.893*
	(0.000)	(0.005)	(0.000)	(0.001)
Market Revenue	-0.0500	-0.0500*	-0.0300	-0.0300
	(0.124)	(0.044)	(0.332)	(0.110)
Market Revenue ^{2}	0.00184	0.00184	0.00130	0.00130
	(0.347)	(0.179)	(0.483)	(0.213)
L.Tran			0.333*	0.333*
			(0.000)	(0.000)
Constant	1.509	1.509*	0.669	0.669*
	(0.175)	(0.000)	(0.527)	(0.001)
Observations	754	754	745	745
p-values in parentheses				
="+ p<0.10	* p<0.05"			

Table 13 Robustness Check Market Level Add Lagged DV

	(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 14 Market Level Use Given Quotes(instead of winning quotes)

	(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 Sentiment Score # Avg Sentiment Score	0.105	1.240+
	-0.847	-0.075
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 15 **Robustness Individual with Text-based Dispersion from Mean Distance**

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 deep learning based text-mining models on business settings in the operations literature. We demonstrate the versatility of word-embedding 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)

??

PLEASE INSERT A TABLE THAT REPORTS THE TEST RESULTS

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?

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Avg	-0.532	-0.532	-0.655	-0.655
	(0.270)	(0.265)	(0.345)	(0.283)
Avg # Avg	0.0448	0.0448	0.0471	0.0471
	(0.543)	(0.530)	(0.660)	(0.614)
Reviews Count	0.0485*	0.0485*	0.0480*	0.0480*
	(0.000)	(0.000)	(0.000)	(0.000)
Avg(Others)	0.00612	0.00612	-0.0285	-0.0285
	(0.958)	(0.967)	(0.873)	(0.876)
Ent Others	1.396*	1.396*	1.762*	1.762*
	(0.018)	(0.021)	(0.020)	(0.012)
Experience	0.133+	0.133+	0.212*	0.212*
	(0.086)	(0.075)	(0.021)	(0.018)
Price Diff	0.206	0.206	0.0861	0.0861
	(0.182)	(0.118)	(0.569)	(0.557)
Market Revenue	-0.00619	-0.00619	-0.0169+	-0.0169*
	(0.424)	(0.330)	(0.074)	(0.023)
Ent Others # Ent Others	-2.252*	-2.252*	-2.626*	-2.626*
	(0.000)	(0.007)	(0.001)	(0.005)
Constant	3.745*	3.745*	4.062*	4.062*
	(0.000)	(0.001)	(0.004)	(0.004)
Observations	3465	3465	4190	4190
p-values in parentheses				
="+ p<0.10	* p<0.05"			

Table 16 Robustness Check Excluding Inactive Installers

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?

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).

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)

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.

NREL (????) Solar photovoltaic technology basics. URL `https://www.nrel.gov/research/re-photovoltaics.html`.

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.