

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

Customer Reviews in an Online 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 investigate 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

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.

3. 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.

3.1. 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

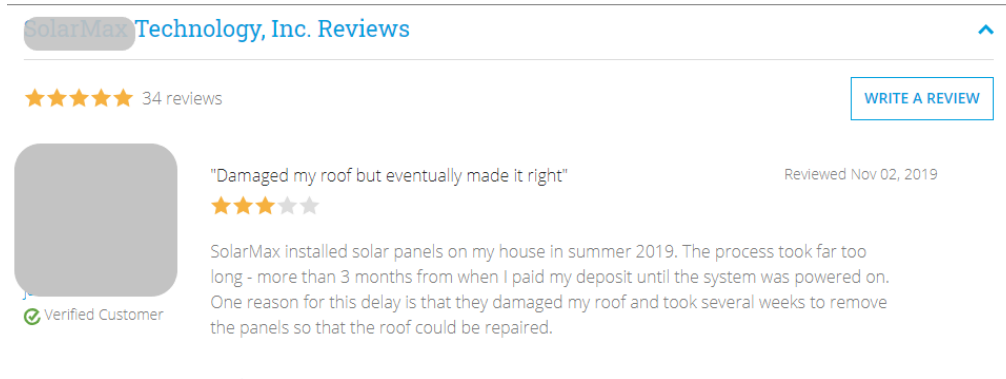


Figure 1 A sample customer review on the marketplace.

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 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 conversations with the online marketplace, the marketplace actively reaches out to solar installers to recruit them to join to platform and help

	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

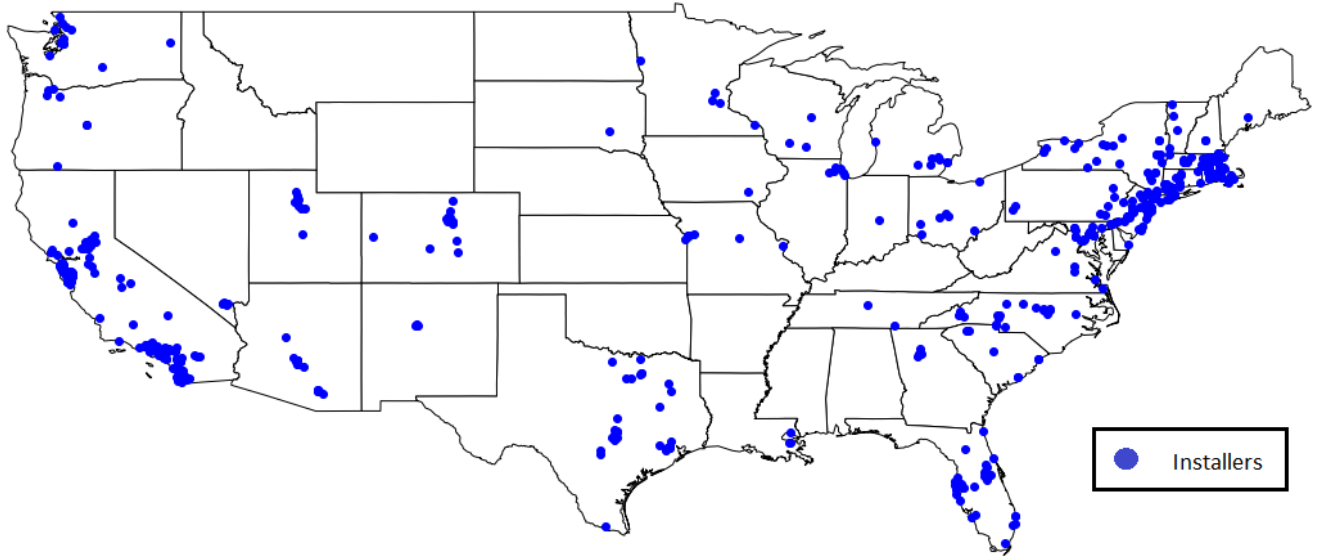


Figure 2 Installers in our data set

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.

3.2. 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 operates only 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 (i.e., ordering points to identify the clustering structure) to identify local markets.

The OPTICS 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 data mining (Breunig et al. 2000), 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, for every installer in the marketplace, we determine the coordinates of its location using its 5-digit zipcode. (Location data are displayed in Figure 2.) To do so, we used the representative coordinates of that zipcode based on the data provided by the US Census (Bureau 2019).

OPTICS The OPTICS algorithm, short for *Ordering Points To Identify the Clustering Structure*, is what we use to cluster installers' coordinates. The OPTICS routine is completed with the following parameter considerations:

min samples: The number of samples in a neighborhood for a point to be considered as a core point. We use 2 as the default value. **metric:** haversine distance. Although not ideal, it better reflected the distance between two Latitude/Longitude points and is still fast enough in the clustering algorithm. **max eps:** The maximum distance between two samples for one to be considered as in the neighborhood of the other. According to the survey that the marketplace conducted, 90.6 percent of customers get a quote from an installer within 100 miles of their property (81.7 percent from 50 miles). (Marsh 2019) We used 90 as the parameter to balance between having enough clusters to make use of the inherent variations and make sure each cluster captured the local market condition. We use this cluster to define our market boundary geographically. 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*.”

[WAS HERE]

4. Model and Measures

In this section we describe our empirical methods. First we introduce the key dependent variable - dispersion of the reviews. We constructed two sets of variables to measure the dispersion of the reviews with both the numeric rating and the reviews text. We first quantify the impact of

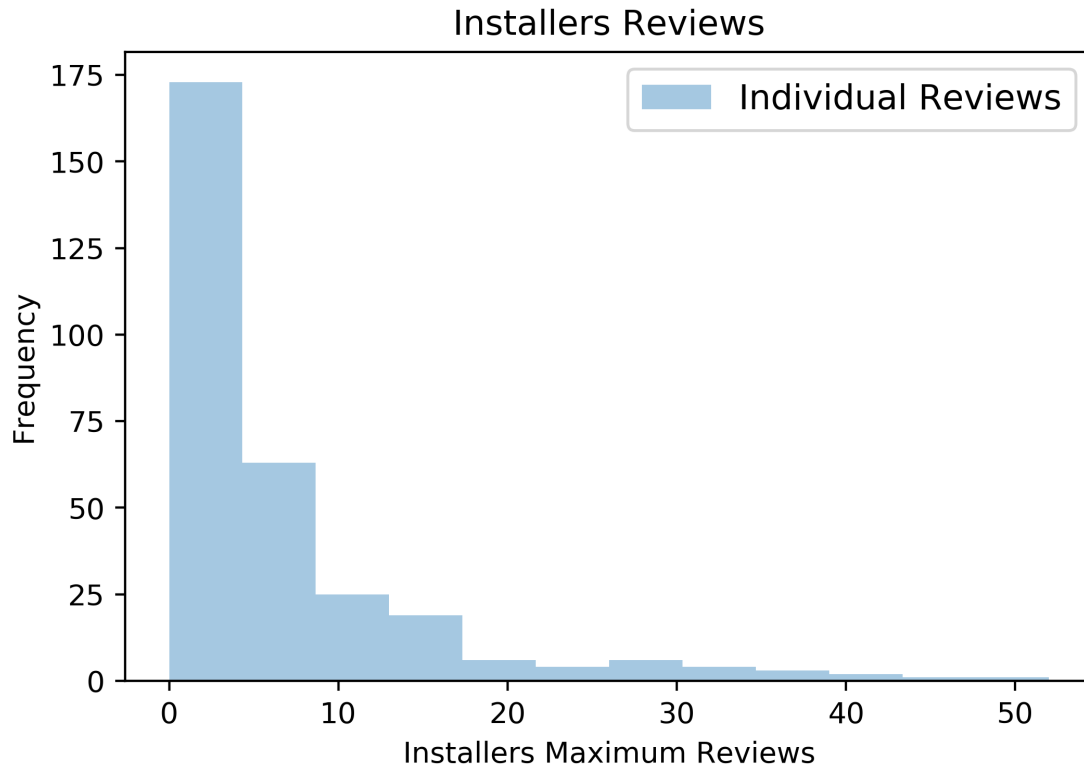


Figure 3 Reviews Histogram by Installer

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

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.

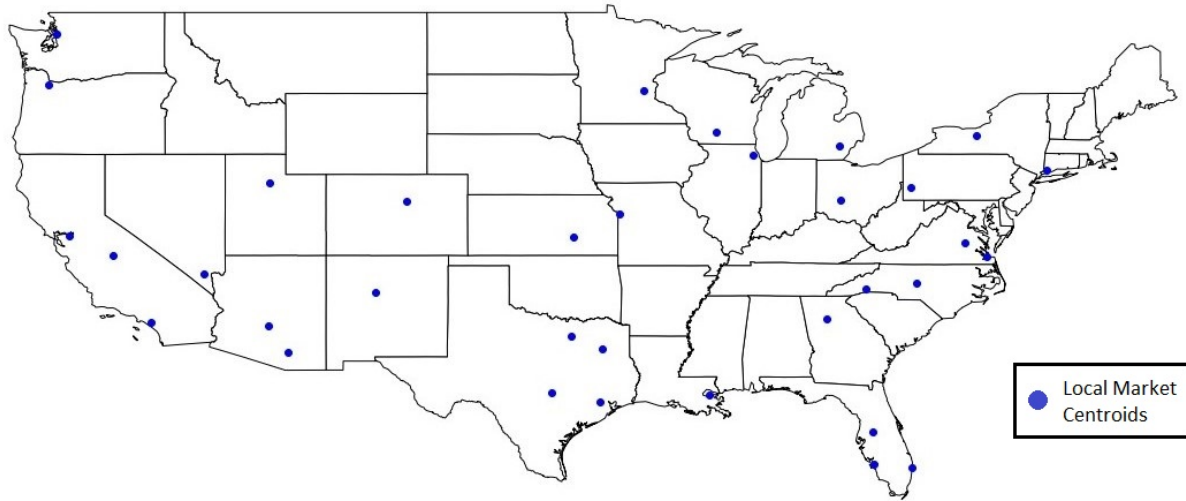


Figure 4 Local Market Centroids

4.1. Measure Reviews Dispersion

We are first interested in measuring the dispersion in ratings, i.e, the quantitative information provided by the ratings. Customer leave a piece of reviews after their solar installation experience. The reviews is composed of a rating (from 1 to 5 stars) and a piece of texts. We first discribe how we measure reviews dispersion from the quantitative ratings information. We also introduce an innovative word embedding model that measures reviews dispersion from the texts data.

4.1.1. Capture Dispersion in Numeric Ratings We use the concept of entropy to capture dispersion in ratings and to construct the independent variable of interests to capture the dispersion of ratings. Entropy is a common measure in information theory (ADD ONE MORE sentence to explain). It can be applied to a collection of a set of discrete probabilistic outcomes, which is our case is the discrete number of ratings (from 1 to 5) that each piece of reviews receive. The formula of computing entropy on a set of discrete values is:

$$H(X) = - \sum P(X) \log(1/P(x)) \quad (1)$$

For example, we want to measure the ratings dispersion on a set of 5 reviews, they all got 4-stars (out of 5). Therefore, we will be applying the calculation on a set $\{4, 4, 4, 4, 4\}$. Following the formula, this set of reviews has an entropy of 0. Alternatively, if we have a set of reviews as $\{3, 5, 3, 5, 4\}$, the entropy of this set of reviews is 1.0549. Although both have the same average rating (4), the second set of ratings are more informative with a higher dispersion, hence has a higher entropy measure.

We apply the entropy calculation on the dataset. For every installer-month, we calculate the ratings entropy on three scopes:

1. Entropy on own reviews, denote as $ENT_{self,i,t}$. This is calculated on the set of reviews that are associated with the focal installer i up to month t .
2. Entropy on peer installers' reviews, up to that month, denote as $ENT_{others,i,t}$. This is calculated on the set of reviews that are associated with all the other installers on focal installer's local market, per market boundaries that were set following steps described in 3.2.

We also calculate entropy on the local market level:

3. Entropy of all reviews on a local market m , up to that month t , which we later denote as $ENT_{mkt,m,t}$. This is calculated on a market-month level data set with market defined in 3.2.

4.2. Installer Level Analysis

We aim to find the connection between ratings dispersion on installer activity intensities. The available data were used to construct a panel data set where the unit of analysis is the the measure of activity intensities ($\log(\text{proposals generated} + 1)$) of a particular installer at a specific local market during a month. We use a regression model with fixed effects and clustered the standard errors on the local market level (individual level yield similar results). We describe our regression models next.

Using the indexes i for installer, m for local market, and t for month, the following regression equation is used to estimate the impact of ratings dispersion (own ratings dispersion: $Ent_{i,self,t}$; others ratings dispersion: $Ent_{i,others,t}$) on focal installer's activity intensities $ActInt_{i,m,t}$:

$$ActInt_{i,m,t+1} = \beta_0 + \beta_{11}Ent_{i,m,others,t} + \beta_2Ent_{i,m,others,t}^2 + Controls_{it} + \alpha_i + \epsilon_{i,m,t} \quad (2)$$

$$ActInt_{i,m,t+1} = \beta_3 + \beta_4Ent_{i,self,t} + \beta_5Ent_{i,self,t}^2 + Controls_{it} + \alpha_i + \epsilon_{i,m,t} \quad (3)$$

$$ActInt_{im,t+1} = \beta_6 + \beta_7Ent_{i,self,t} + \beta_8Ent_{i,self,t}^2 + \beta_9Ent_{im,others} + \beta_{10}Ent_{im,self,t}^2 + Controls_{it} + \alpha_i + \epsilon_{imt} \quad (4)$$

The error term ϵ represents factors that affect installer activity intensity that are unobservables in the data. α_i represents the installer level fixed effects. $Controls_{it}$ are all other installer-level or market level control variables we include in order to capture factors that are irrelevant to reviews dispersion. We detail our selection of control variables next.

4.2.1. Control Variables

State: State dummies are included to account for state level policy effects. (ELABORATE)

Price: Price is an important factor. Although we do not model installers' pricing strategy, we want to control for the impact of price on installers' activities. We decide to look at the price difference between the focal installer and the others. We use Tracking the Sun data to find the installers' prices via matching name and Zipcode. We also use the unit price: price per KW. Price per KW is a common way to assess the price level of a solar system, as the final price tag of the solar system will be dependent on the size. We then compute the variable $PriceDiff_{i,t}$ as the difference in unit price between installer and the average unit price of their competitors on the local market.

Average Ratings: The average ratings of installer themselves $avg_{i,t}$ and the average ratings of their competitors $avg_{others,t}$ on the market.

Experience: The number of years the installer has been installing solar systems. We obtain that information from installers' website.

Local Markets Condition: Once the algorithm gave us the clusters that defines market divisions, we augment the data with Tracking The Sun data to capture local market conditions. We use the same Market definition from 3.2 and computed the sum of all solar installations within a market during a month, denote as $MarketRev_{mt}$. The market revenue variable measures the total opportunities of solar installations on that market.

In addition, installer level fixed-effects are included to control for time-invariant characteristics of each installers.

4.3. 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.

$$MarketActivity_{m,t+1} = \beta Ent_{m,t} + \beta Ent_{m,t}^2 + Controls + \epsilon_{mt} \quad (5)$$

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.

4.3.1. 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

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

5. Results

We now present the results of our analysis. We first illustrate the important link between Ratings Dispersion and Installer activity level with a simple matching framework. We then present the summary statistics, followed by results of the panel regression models that focus on the link between ratings dispersion and activities intensity.

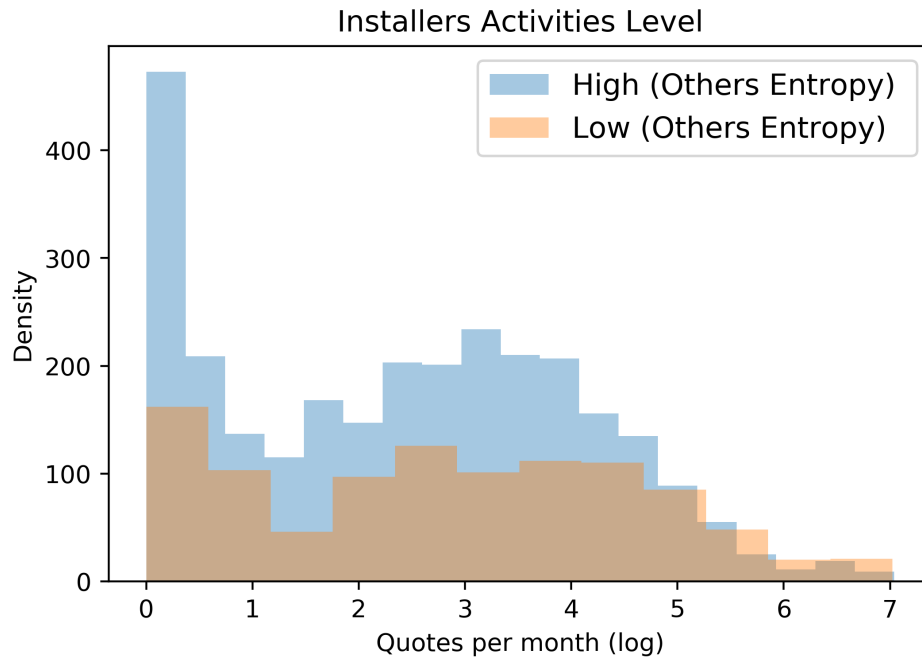


Figure 5 Histogram High vs. Low Entropy

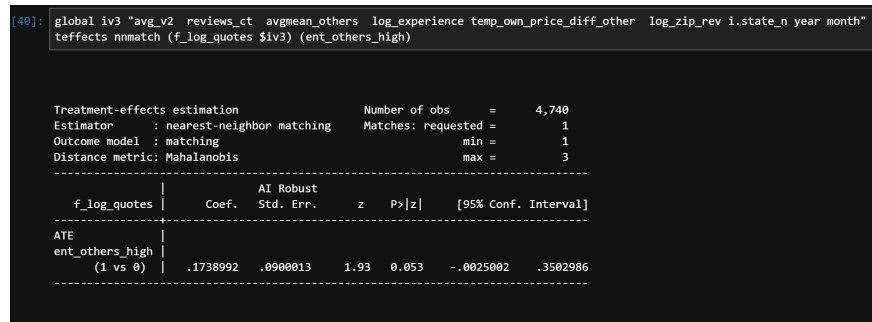


Figure 6 (TEMPORARY: Matching illustrate the impact of high Entropy)

5.1. Summary Statistics

For the individual level analysis data, in table 3 and 5 we present the descriptive statistics and correlation matrix. Likewise table 4 and 6 contains that on the local market level. We find that the correlations are generally in the expected direction and not a huge concern for the validity of regression analysis.

The *Entropy* measure: We found that although the activity levels vary, one key factor that would explain the variation is the local market ratings *Entropy*. To illustrate this point, we divide the observations into two groups based on $ENT_{i,others,t}$ level (their peers ratings dispersion/ (others' ratings entropy)). With a median level equals 0.27, the ENT_{others} level higher than 0.27 were coded as 'high ratings dispersion', or 'ent-others-high'. A matching model matches installers on

	Obs	Mean	SD	Min	Max
Ratings Average	4953	4.880346	.3971615	1	5
Reviews Count	8113	3.397387	5.80235	0	52
Entropy Own Reviews	5340	.0981426	.2199734	0	1.209574
Entropy Others Reviews	5340	.2247373	.1915538	0	1.070593
Experience	8113	5.693455	5.748559	-4	43
Price Diff	8113	-.0104862	.4354344	-2.171179	6.641788
Market Rev	8113	3.59e+07	1.17e+08	0	4.85e+09

Table 3 Summary Statistics Individual Level

VARIABLES	N	mean	min	max
Entropy	804	0.176	0	1.055
Mkt Revenue	1,528	8.202	0	22.30
Rating Avg	1,027	4.888	3	5
Experience(Avg)	1,528	4.870	0	27
Price Diff(Avg)	1,528	0.0160	-1.106	1.970
Transaction	1,528	0.624	0	4.522
Activity	1,528	3.048	0	8.362

Table 4 Summary Statistics Market 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 5 Correlation Individual Level

every important variables (own ratings, others' average ratings, experience, price difference, local market condition, state) except for Ent_{others} . Matching suggest that a higher ratings entropy is associated with a significant higher level activity intensity, as presented in 6 and histogram in 5. In the matching result presented in figure 6, the binary variable representing 'ent-others-high' is significant with p-value 0.05 and coefficient 0.17. We investigate the activity levels differences further with a panel regression model in the next section.

5.2. Individual Installer

We first present the results pertaining to the impact of reviews entropy on individual installers as estimated by the regression models. These results are presented in table 8 and 9.

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 6 Correlation Market Level

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Avg	-0.506 (0.476)	-0.506 (0.432)	-0.294 (0.651)	-0.294 (0.652)
Avg(Others)	-0.0615 (0.718)	-0.0615 (0.685)	-0.104 (0.545)	-0.104 (0.485)
Avg # Avg	0.0482 (0.668)	0.0482 (0.623)	0.0156 (0.877)	0.0156 (0.872)
Reviews Count	0.0541* (0.000)	0.0541* (0.000)	0.0547* (0.000)	0.0547* (0.000)
Experience	0.310* (0.007)	0.310* (0.000)	0.294* (0.000)	0.294* (0.000)
Price Diff	-0.0222 (0.796)	-0.0222 (0.828)	-0.0432 (0.577)	-0.0432 (0.635)
Market Revenue	-0.0142+ (0.053)	-0.0142+ (0.059)	-0.0149* (0.044)	-0.0149* (0.033)
Constant	3.368* (0.019)	3.368* (0.010)	4.032* (0.003)	4.032* (0.003)
Observations	5772	5772	5772	5772
p-values in parentheses				
= "+" p<0.10		* p<0.05"		

Table 7 Individual Level Regression with Only Ratings Average

Signal : the direct impact from ‘Signal’, or the Average ratings, are negative albeit not consistently significant, after we controlled for installer and local market level characteristics.

Noise : we found consistent results associated with the set of variables representing ‘noise’. The estimates suggest that the direct effect of ratings dispersion (β_{e1} in equation X) on activity intensity is positive and statistically significant, and the second order effect (β_{e2} in equation X) is negative and statistically significant. In order words, the regression estimates indicates that for individual installers dispersion (of others’ ratings) increase activity when dispersion is small, but deters activity when dispersion is large.

We plot the effects in Figure 8 to further illustrate the non-linear effect of entropy on activity intensity. We use the estimated regression coefficient from the model in table 8 to generate the

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Entropy Others	1.767*	1.767*	1.921*	1.921*
	(0.018)	(0.011)	(0.005)	(0.003)
Entropy Others ²	-2.635*	-2.635*	-2.710*	-2.710*
	(0.001)	(0.005)	(0.000)	(0.002)
Avg	-0.679	-0.679	-0.360	-0.360
	(0.317)	(0.258)	(0.576)	(0.572)
Avg(Others)	-0.0323	-0.0323	-0.0489	-0.0489
	(0.856)	(0.859)	(0.797)	(0.784)
Avg # Avg	0.0513	0.0513	0.00758	0.00758
	(0.624)	(0.575)	(0.938)	(0.935)
Reviews Count	0.0479*	0.0479*	0.0492*	0.0492*
	(0.000)	(0.000)	(0.000)	(0.000)
Experience	0.214*	0.214*	0.193*	0.193*
	(0.020)	(0.016)	(0.005)	(0.008)
Price Diff	0.0866	0.0866	0.00151	0.00151
	(0.566)	(0.554)	(0.991)	(0.991)
Market Revenue	-0.0170+	-0.0170*	-0.0172+	-0.0172*
	(0.069)	(0.022)	(0.062)	(0.014)
Constant	4.096*	4.096*	4.426*	4.426*
	(0.004)	(0.004)	(0.002)	(0.004)
Observations	4200	4200	4200	4200
p-values in parentheses				
=" + p<0.10 * p<0.05"				

Table 8 Individual Level with Entropy of Others' Ratings

marginal effects. As is apparent from the margins plot, the activity intensity first rise then fall with the ratings dispersion.

5.3. Local Markets

We now move to discuss the ratings dispersion on total transactions on local market level. The results are presented in table 11. 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.

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 10.

	(1)	(2)	(3)	(4)
	F.Activity	F.Activity	F.Activity	F.Activity
Entropy Others	1.748*	1.748*	1.918*	1.918*
	(0.020)	(0.010)	(0.005)	(0.002)
Entropy Others ²	-2.686*	-2.686*	-2.767*	-2.767*
	(0.001)	(0.003)	(0.000)	(0.001)
Entropy Own	2.680*	2.680*	2.445*	2.445*
	(0.004)	(0.005)	(0.001)	(0.004)
Entropy Own ²	-3.303*	-3.303*	-2.908*	-2.908*
	(0.020)	(0.015)	(0.012)	(0.017)
Avg	-1.177*	-1.177*	-0.898	-0.898
	(0.050)	(0.035)	(0.120)	(0.133)
Avg(Others)	-0.0310	-0.0310	-0.0446	-0.0446
	(0.871)	(0.867)	(0.823)	(0.805)
Avg # Avg	0.132	0.132	0.0963	0.0963
	(0.146)	(0.113)	(0.266)	(0.263)
Reviews Count	0.0423*	0.0423*	0.0439*	0.0439*
	(0.000)	(0.000)	(0.000)	(0.000)
Experience	0.210*	0.210*	0.190*	0.190*
	(0.023)	(0.016)	(0.005)	(0.008)
Price Diff	0.0932	0.0932	0.0138	0.0138
	(0.514)	(0.510)	(0.914)	(0.918)
Market Revenue	-0.0162	-0.0162*	-0.0166+	-0.0166*
	(0.104)	(0.033)	(0.086)	(0.019)
Constant	4.549*	4.549*	4.900*	4.900*
	(0.002)	(0.002)	(0.001)	(0.001)
Observations	4200	4200	4200	4200
p-values in parentheses				
=”+ p<0.10	* p<0.05”			

Table 9 Individual Level with Entropy of Others' Ratings and Own Ratings

	(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 10 Market Level with Only Ratings Average

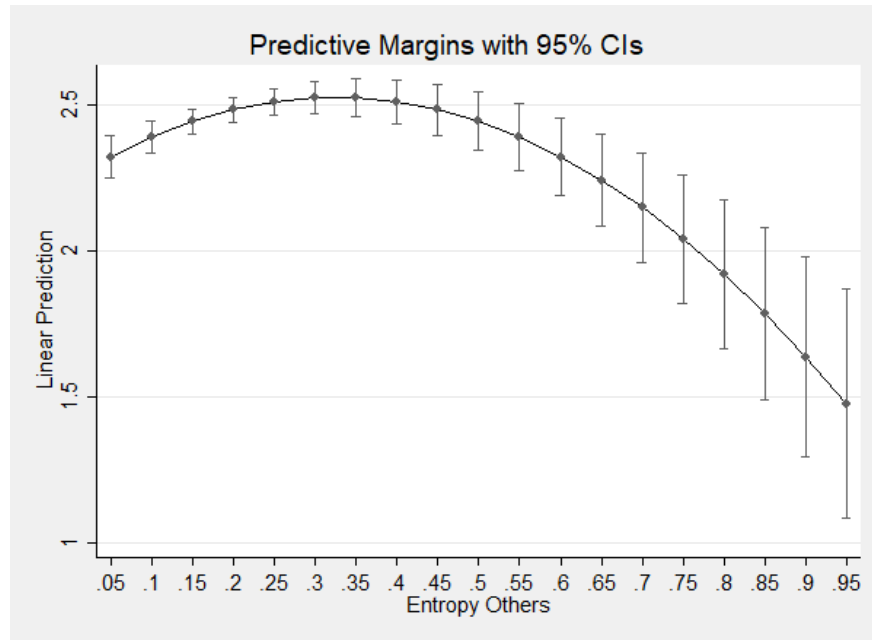


Figure 7 Marginal Impact of Entropy of Reviews on Individual Level Activity

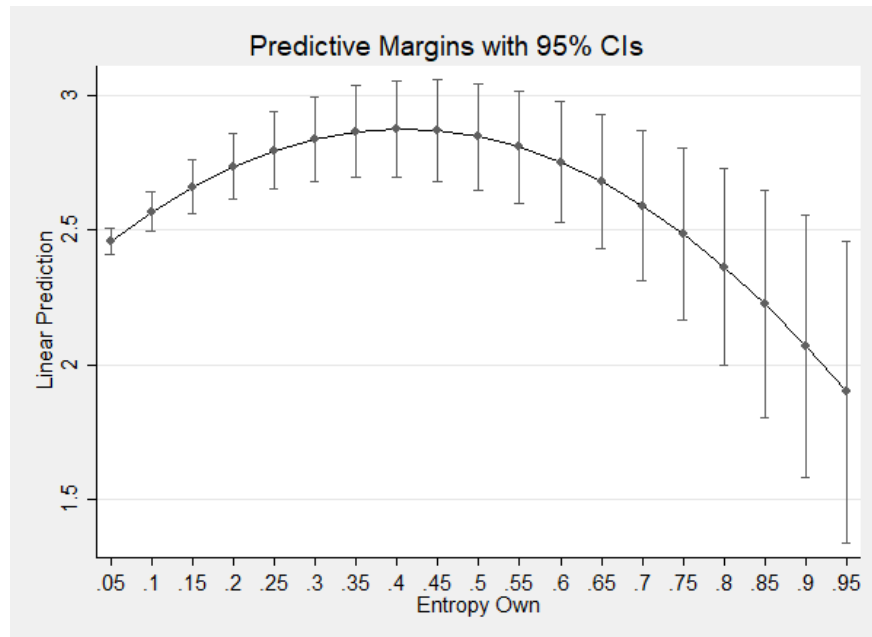


Figure 8 Marginal Impact of Entropy of Reviews on Individual Level Activity

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.

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

Table 11 Market Level Use Winning Quotes and Quantitative Dispersion Measures

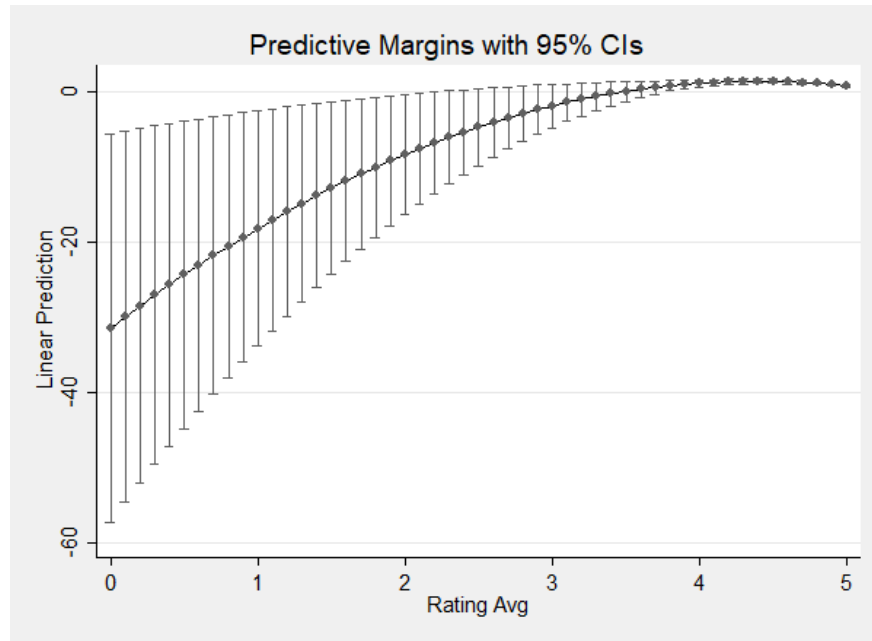


Figure 9 Marginal Impact of Average Ratings on Market Level Activity

6. 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

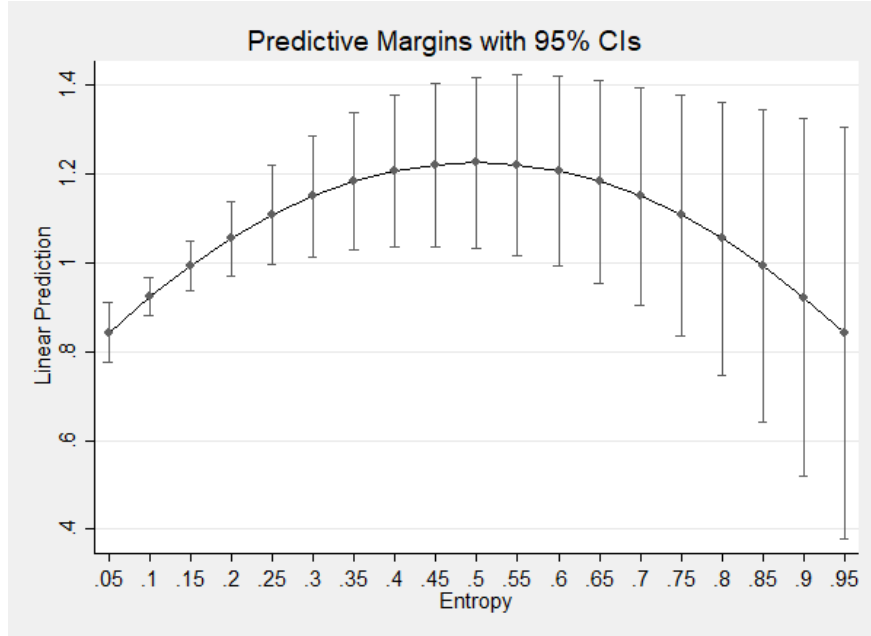


Figure 10 Marginal Impact of Market Reviews Entropy of Reviews on Market Level Activity

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.

6.1. 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 significantly ($corr = 0.8239$). It is further illustrated in figure 11 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.

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

Table 12 Individual Level Use Average Rating and Text-based dispersion

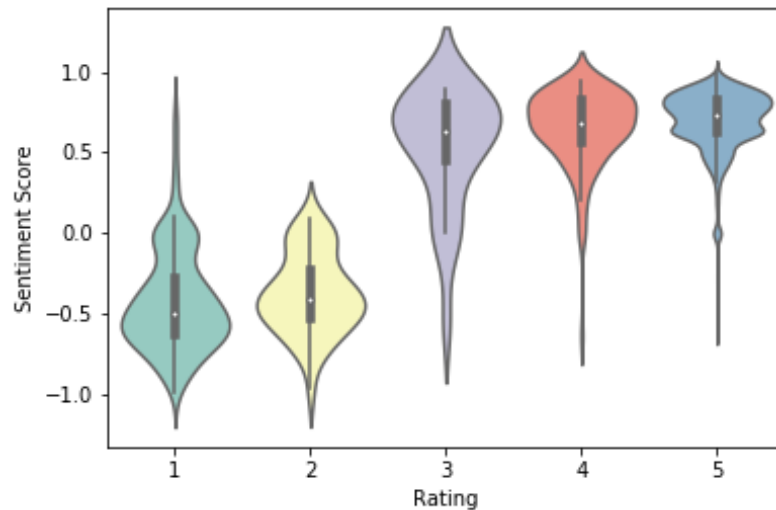


Figure 11 Sentiment Scores and Ratings

	(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 13 Individual Analysis Use Sentiment Score and Entropy

6.2. 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.

6.2.1. 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 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.

	(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 14 Use Sentiment score and Text-based Entropy

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 the two vectors is calculated as follows:

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

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

	(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 15 Market Level Use Use Sentiment Score and Text-based Dispersion

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 Bidirectional 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 appearances 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 examples 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’ vectorization 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 768×1 , 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 $ENT_{self,i,t}, ENT_{Others,i,t}, ENT_{m,t}$

6.3. 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 13 and table 14 . Likewise we run the same regression on the market level data and presented the results in table 15. 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

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

$$ActInt_{i,m,t+1} = \delta pro_{i,t} + \beta_1 Ent_{i,m,others,t} + \beta_2 Ent_{i,m,others,t}^2 + controls + \epsilon_{i,m,t} \quad (6)$$

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 β (CITE ECONOMETRIC stuff).

7.2. Robustness with different local market division

Although many similar studies used ZIP code to define 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:

$$ActInt_{i,m,t+1} = \gamma ActInt_{i,m,t-1} + Ent_{i,m,others,t} + Ent_{i,m,others,t}^2 + controls + \epsilon_{i,m,t} \quad (7)$$

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

Table 16 Robustness Check Add Lagged DV

$$ActInt_{im,t+1} = \gamma Ent_{im,t-1} + \beta_3 + \beta_4 Ent_{i,self,t} + \beta_5 Ent_{i,self,t}^2 + Controls + \epsilon_{imt} \quad (8)$$

$$ActInt_{im,t+1} = \gamma Ent_{im,t-1} + \beta_6 + \beta_7 Ent_{i,self,t} + \beta_8 Ent_{i,self,t}^2 + \beta_9 Ent_{im,others} + \beta_{10} Ent_{im,self,t}^2 + Controls + \epsilon_{imt} \quad (9)$$

We expect γ estimates to be positive. The results are still consistent as the β coefficients associated with $Ent_{others}(Ent_{others}^2)$ are still positive (negative) as presented in table 16 and 17.

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

$$MarketActivity_{m,t+1} = \gamma MarketActivity_{m,t-1} + \beta Ent_{m,t} + \beta Ent_{m,t}^2 + Controls + \epsilon_{mt} \quad (10)$$

and the results, presented in table 18, are still consistent.

	(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 17 Robustness Check with Lagged Variable and Own Entropy

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 19 shows.

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 20.

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 21. The first two columns

	(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 18 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 19 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 20 Robustness Individual with Text-based Dispersion from Mean Distance

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:

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 13 and 14 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.

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

Table 21 Robustness Check Excluding Inactive Installers

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 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

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 INSTALLER AND HISTOGRAM? 2) NUMBER OF INSTALLERS IN EACH STATE FOR TOP 10 STATES

QUESTIONS: 1) HOW MANY OF THE INSTALLERS DO NOT HAVE A REVIEW? 2) WHAT IS THE PERCENTAGE OF THOSE INSTALLERS IN THE LOCAL MARKET WIN AND SUBMITTED PROPOSALS? 3) WHAT DO YOU ASSUME 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 entotners 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/>.
- Breunig MM, Kriegel HP, Sander J (2000) Fast hierarchical clustering based on compressed data and optics. *European Conference on Principles of Data Mining and Knowledge Discovery*, 232–242 (Springer).
- Bureau UC (2019) Gazetteer files. URL <https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.2018.html>.
- 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.
- Marsh J (2019) Solar providers near me: How to find them in 2020: Energysage. URL <https://news.energysage.com/best-way-find-solar-panel-installers-near-me/>.
- 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.
- 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.