

Herbst, Chris M. et al.

Working Paper

What Do Parents Value in a Child Care Provider? Evidence from Yelp Consumer Reviews

IZA Discussion Papers, No. 11741

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Herbst, Chris M. et al. (2018) : What Do Parents Value in a Child Care Provider? Evidence from Yelp Consumer Reviews, IZA Discussion Papers, No. 11741, Institute of Labor Economics (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/185201>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

DISCUSSION PAPER SERIES

IZA DP No. 11741

**What Do Parents Value in a Child Care
Provider? Evidence from Yelp Consumer
Reviews**

Chris M. Herbst
Kevin C. Desouza
Saud Alashri
Srinivasa Srivatsav Kandala
Mayank Khullar
Vikash Bajaj

AUGUST 2018

DISCUSSION PAPER SERIES

IZA DP No. 11741

What Do Parents Value in a Child Care Provider? Evidence from Yelp Consumer Reviews

Chris M. Herbst

Arizona State University and IZA

Kevin C. Desouza

Queensland University of Technology

Saud Alashri

King Abdulaziz City for Science and Technology

Srinivasa Srivatsav Kandala

Arizona State University

Mayank Khullar

Microsoft Corporation

Vikash Bajaj

Arizona State University

AUGUST 2018

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

What Do Parents Value in a Child Care Provider? Evidence from Yelp Consumer Reviews¹

This paper exploits novel data and empirical methods to examine parental preferences for child care. Specifically, we analyze consumer reviews of child care businesses posted on the website Yelp.com. A key advantage of Yelp is that it contains a large volume of unstructured information about a broad set of child care programs located in demographically and economically diverse communities. Thus our analysis relies on a combination of theory- and data-driven methodologies to organize and classify the characteristics of child care that are assessed by parents. We also use natural language processing techniques to examine the affect and psychological tones expressed in the reviews. Our main results are threefold. First, we find that consumers overall are highly satisfied with their child care provider, although those in higher-income markets are substantially more satisfied than their counterparts in lower-income markets. Second, the program characteristics most commonly evaluated by consumers relate to safety, quality of the learning environment, and child-teacher interactions. However, lower- and higher-income consumers evaluate different characteristics in their reviews. The former is more likely to comment on a program's practical features, such as its pricing and accessibility, while the latter is more likely to focus on the learning environment. Finally, we find that consumers in lower-income markets are more likely to display negative psychological tones such as anxiety and anger in their reviews, particularly when discussing the nature of their interactions with program managers and their child's interactions with teachers.

JEL Classification: J13

Keywords: child care, early childhood education, preferences, Yelp, content analysis, machine learning

Corresponding author:

Chris M. Herbst
School of Public Affairs
Arizona State University
411 N. Central Ave., Suite 420
Phoenix, AZ 85004-0687
USA
E-mail: chris.herbst@asu.edu

¹ We gratefully acknowledge financial support for this project provided by the College of Public Service and Community Solutions at Arizona State University.

I. Introduction

Over the past few decades, a scholarly consensus has coalesced around the idea that high-quality child care is important to children’s short-run development and long-run schooling and labor market success (Auger et al., 2014; Gathman & Sass, 2017; Havnes & Mogstad, 2011; Herbst, 2013; 2017; Keys et al., 2013; Rossin-Slater & Wust, 2016). Yet many U.S. children—particularly those in low-income families—attend child care programs that are of low- to mediocre-quality (National Research Council & Institute of Medicine, 2000; NICHD ECCRN, 2005). In addition, the child care workforce is plagued by low skills, low and stagnant wages, and high turnover (Bassok et al., 2013; Herbst, 2018; Boyd-Swan & Herbst, 2018). As a result, policymakers have come under increasing pressure to adopt measures that improve the quality of children’s early care and education experiences.

A potential explanation for the low level of quality points to constraints on the parent-side of the market. Indeed, it is argued that parents possess imperfect information about what constitutes “high-quality” child care, and they undervalue the positive external benefits from consuming such care. Information problems are likely to arise in the child care market because caregiving is an experience good—one whose key quality features are not easily observed by consumers prior to purchasing it. When parents cannot make informed decisions, child care providers have an incentive to produce low-quality services and offer those services at lower prices than would be the case if parents were perfectly informed. As a result, high-quality providers are forced out of the market—or do not enter the market in the first place—leaving those willing to offer mediocre services (Blau, 2001).

This line of reasoning seems broadly consistent with the existing literature on parents’ child care decisions. One stream of research, which studies the child care search process, shows that parents generally allocate little time to the search, consider only a small number of options, and rely primarily on friends and family for recommendations (NSECE, 2014). Another line of work attempts to tease out parents’ revealed preferences for specific quality and non-quality features of child care (Forry et al., 2013a). This work shows that although parents claim to value high-quality, education-focused programs, actual decisions are

equally—if not more—driven by such practical considerations as program location, cost, and convenience (Chaudry et al., 2011; Mamedova et al., 2013; NSECE, 2014). A final body of work attempts to assess directly whether parents are accurate evaluators of quality by comparing their ratings of specific program features with those of trained observers (Cryer & Burchinal, 1997; Cryer et al., 2002; Mocan, 2007). This work consistently shows large disagreements between the ratings of parents and the observers, with former rating the quality of their child’s program more favorably than the latter.

While the evidence seems consistent with the notion that parents are not able to accurately assess the characteristics and quality of their child care options, prior work is limited in several respects. First, the data collection strategies used to study parent preferences range from large nationally representative surveys (e.g., Mamedova & Redford, 2013) and single-site surveys (e.g., Bassok et al., 2018a) to small-scale focus groups (e.g., Forry et al., 2013b) and one-on-one interviews (e.g., Barbarin et al., 2006). While the results from surveys are more likely to be generalizable, they typically include only a small number of close-ended, highly-structured questions about a narrow set of child care characteristics. Conversely, although the interview-based studies provide textured information about the decision-making process, the results are less generalizable. Furthermore, the small sample sizes in these studies preclude an analysis of how parent preferences vary with education or income levels. Second, many studies are deductive investigations of parent preferences, in that the survey instruments are designed with strong *a priori* beliefs about the characteristics of child care that constitute “high-quality” (e.g., Cryer et al., 2002). Therefore, these instruments generally do not include items that inquire about cost, location, and availability—all of which are found to be crucial to parents’ child care decisions when included in the choice set (e.g., Bassok et al., 2018b; Chaudry et al., 2011). A final limitation is that the mode of data collection for many studies—telephone surveys or face-to-face interviews—may engender a type of social desirability bias, such that parents feel pressured to report higher levels of satisfaction with their child care arrangement, or stronger preferences for its quality-related attributes, than is actually the case. In fact, such biases may explain the

disconnect between parents' high levels of satisfaction with their provider and the comparatively low quality ratings that such programs receive from trained observers.

In this paper, we exploit novel data and empirical methods to provide new evidence on parents' satisfaction with and evaluation of their child care program. We construct a dataset of consumer reviews of child care businesses in the 40 largest U.S. cities from the website Yelp.com. Founded in 2004, Yelp is the predominant online platform for hosting consumer ratings and narrative reviews of virtually all business-types. Several attributes make Yelp a powerful laboratory for studying parental assessments of child care. First, it contains a rapidly growing number of child care reviews. Between 2005 and 2017, the number of annual reviews increased from 43 to over 15,000 in our 40 cities. Altogether our dataset includes nearly 50,000 unique reviews of about 9,800 businesses. Furthermore, our dataset—while perhaps not fully representative of all child care programs in these cities—is large enough to examine variation in consumer preferences by characteristics of the local market (e.g., household income). Second, Yelp users provide unstructured, narrative reviews of child care programs. These attributes are advantageous because consumers have the autonomy to evaluate any dimension of the program deemed important, and to do so using a variety of emotional tones. Our analysis therefore relies on a combination of theory- and data-driven methodologies to organize and classify the characteristics of child care that are evaluated by parents. We also use a natural language processing tool to examine the affect and emotions expressed in the reviews. Third, Yelp allows users to vote on the usefulness of each review. Under the assumption that those casting useful votes are more likely to be searching for child care, this information is valuable for examining the preferences of *potential* rather than *actual* consumers.

Our analysis begins by using a combination of deductive and inductive methodologies to study parent assessments of child care. We first rely on the Early Childhood Environment Rating Scale-Revised (ECERS-R), a widely used measure of classroom quality, as a framework for organizing the type and frequency of quality-related language used in the text-based reviews (Harms et al., 1998). The ECERS-R measures multiple dimensions of the child care environment—including space and furnishings, language

and reasoning, activities, and interactions—and we code whether parents’ evaluation of their child care program overlaps with the content in each ECERS-R domain. Thus by cataloguing reviews according to the ECERS-R domains, this paper sheds light on whether parent assessments are aligned with the tools used by researchers to measure quality. We then utilize Latent Dirichlet Allocation (LDA), an unsupervised machine learning algorithm, to locate additional characteristics of child care discussed in the Yelp reviews. Finally, given that parents identify several practical features of child care as being important, we code whether the reviews discuss program costs, availability, and religious affiliation. Altogether our analysis examines parent evaluations of 14 distinct child care characteristics. The next step of our analysis applies a natural language processing tool called Linguistic Inquiry and Word Count (LIWC) to classify the affect and emotional tones present in the reviews. For example, we code whether parents express positive or negative emotions, anxiety, anger, and a variety of other psychological states. Importantly we are able to determine whether the individual sentences describing the 14 child care domains express positive or negative sentiments. The final set of analyses examine the number of useful votes received by reviews that comment on each child care characteristic. A major objective of our paper is to examine whether parent assessments of child care vary across poorer and wealthier markets. Therefore, in all analyses we examine heterogeneity in the consumer reviews over the distribution of local household income.

Our primary findings are summarized as follows. First, we find that consumers overall are highly satisfied with their child care provider: the average Yelp rating is 4.3 stars (out of five), with 76 percent of consumers giving their program five stars. Nevertheless, there is substantial variation in ratings over the distribution of income. For example, 85 percent of consumers located in the top decile of county household income give their program a five-star rating, compared to 65 percent of those in the bottom income decile. Second, we show that consumers’ text-based reviews focus heavily on program quality. Indeed, the characteristics most commonly evaluated by consumers relate to child-teacher interactions, the learning and academic environment, and safety. Furthermore, we identify two other salient features of child care that have been overlooked in previous studies: consumer experiences throughout the search

process (including the facility visit) and interactions with program managers and owners. Less frequently discussed are topics related to the cost and availability of care, a result that contrasts with some previous work. Again, however, we find that lower- and higher-income consumers evaluate different characteristics in their reviews. The former is more likely to comment on a program’s practical features, such as its pricing and accessibility, while the latter is more likely to focus on the learning environment. Third, we find that across all 14 program domains, consumers in higher-income markets display more positive affect when describing their child care experiences; lower-income consumers, on the other hand, are more likely to express a range of negative emotions, including anxiety and anger. Finally, Yelp reviews posted in lower-income markets garner substantially more useful votes than those in higher-income markets, especially when the reviews focus on the practical aspects of child care (i.e., cost and availability).

The remainder of the paper proceeds as follows. Section II summarizes the relevant literature on parents’ child care search behavior and preferences. In Section III, we discuss how the introduction of Yelp may influence the demand for child care services. Section IV describes the Yelp dataset and analytic strategy. Our main results are presented in Section V and discussed in Section VI. We conclude the paper in Section VII with a discussion of policy implications.

II. Relevant Literature

Three streams of research are relevant to the current paper. We begin with a discussion of how parents search for child care, including the duration of the search, the number of programs considered, and the sources of information used to make a selection. We then discuss previous work on parents’ priorities and preferences when selecting a program, including the relative importance of various quality and non-quality attributes in the selection process. Finally, we summarize the literature on parents’ satisfaction with and evaluation of the characteristics of their child’s program.

The Child Care Search

The available evidence suggests that parents generally allocate little time to the child care search, they consider a small number of programs before making a decision, and they rely on informal networks

for information on local programs. One nationally representative study of low-income families reports that 41 percent of parents made a decision within one day (Layzer et al., 2007), while another study finds that 82 percent of welfare applicants in Minnesota finished the search within two weeks (Forry et al., 2014). These short search durations are explained, in part, by the fact that parents do not visit many child care providers during the selection process. For example, one study finds that 38 percent of parents overall consider only one arrangement during their search (NSECE, 2014). Among low-income families, the share considering one option appears to be substantially higher (Anderson et al., 2005; Bassok et al., 2018b).

Regarding information sources, nationally representative data reveal that parents overwhelmingly rely on friends and family members for advice and recommendations on child care programs (NSECE, 2014). In contrast, relatively few families seek assistance through such formal channels as state-administered resource and referral lists and Quality Rating and Improvement System (QRIS) websites. Although these search patterns apply to both low- and high-income families, it appears that a reliance on friends and family is more prevalent among wealthy families. For example, one study finds that 68 percent of families at or above 300 percent of the federal poverty line (FPL) seek recommendations from friends and family, while 54 percent of those below the FPL do so (NSECE, 2014).

Parental Preferences for Child Care Characteristics

There is a large literature dedicated to understanding which features of child care are the most salient to parents during the search process (Forry et al., 2013a). At the broadest level, nationally representative surveys find that “reliability” and “learning activities” are the most important features in a program, followed by “cost” and “location” (Mamedova & Redford, 2013). Parents’ dual emphasis on programs’ practical and quality-related features is similarly borne out in studies that ask parents to list or rank-order their preferences. On the one hand, parents value child care programs that promote learning and social-emotional development; those that are staffed by highly educated and experienced teachers who cultivate warm and trusting relationships with children; and those providing nutritious meals (Barbarin et al., 2006; Raikes et al., 2012; Rose & Elicker, 2008; Shlay et al., 2005). However, parents also reveal strong

preferences for programs whose operating hours align with their work schedules, are located close to home or work, and are affordable (Barbarin et al., 2006; Bassok et al., 2018b; Rose & Elicker, 2008).

Important for the purposes of this study is the research examining heterogeneity in parent preferences by education or income level. Generally speaking, household income and maternal education are strongly correlated with child care choices, with high-income/-education families more likely to select formal child care arrangements (Early & Burchinal, 2001; NICHD ECCRN, 1997; Pungello & Kurtz-Costes, 2000). In addition, nationally representative surveys find that disadvantaged families are more likely to rate “location”, “cost”, and “availability” as being important to the child care decision than their advantaged counterparts (Mamedova & Redford, 2013). Such patterns are confirmed in smaller-scale studies asking detailed questions about preferences: disadvantaged parents reveal a stronger interest in the practical features of child care (e.g., safety, convenience, and cost), while advantaged parents are more likely to emphasize its quality features (e.g., curriculum and staff training and education) (Gordon & Hognas, 2006; Johansen et al., 1996; Peyton et al., 2001). Specifically, one study finds that low-skilled parents rate affordability, close distance from home or work, and flexibility in hours of operation as being highly important features, while high-skilled parents rate teachers’ education levels and the curriculum-type as the most features (Rose & Elicker, 2008).

Parental Satisfaction With and Evaluation of Their Child Care Provider

Parents consistently report high levels of satisfaction with their child care arrangement. In a recent nationally representative survey, fully 88 percent of parents with preschool-age children rated their child’s program as “very good” or “excellent” (National Public Radio, 2016). Such high ratings are evident when parents are asked about specific characteristics (e.g., Cryer et al., 2002), and these patterns apply to low-income families, who generally use lower-quality care than economically advantaged families. For example, a study of low-income families in Louisiana finds that 69 percent of parents are very satisfied with their child’s program and would likely choose the program again (Bassok et al., 2018b). Another study of low-

income families receiving child care subsidies finds that approximately three-quarters of mothers rate the quality of their child’s program as either “excellent” or “perfect” (Raikes et al., 2012).

Despite these high levels of satisfaction, the evidence suggests that parents’ preferences are not always borne out in their actual consumption decisions. In particular, several studies uncover a disconnect between the characteristics of child care that are reported to be important to parents—i.e., high levels of quality—and the characteristics of the program actually used—i.e., high levels of convenience and affordability (Chaudry et al., 2011; Sandstrom & Chaudry, 2012). In addition, the evidence suggests that parents are not able to accurately assess the quality of their child’s program. Indeed, several studies compare parents’ quality ratings with those of trained observers on an identical set of scales (e.g., Cryer & Burchinal, 1997; Cryer et al., 2002; Mocan, 2007). This work shows large disagreements between the ratings of parents and observers, with the former rating the quality of their child’s program more favorably than the latter. For example, using a well-established measure of quality, Cryer and Burchinal (1997) show that parents rate their program an average of 6.1 (on a seven-point scale), while the observers give the same programs a rating of 3.5. In addition, one recent study examines whether low-income parents’ satisfaction with their child care provider is associated with a variety of program characteristics (Bassok et al., 2018a). This work finds that parent satisfaction scores are unrelated to most program features, including measures of overall quality, teacher training and education, group size, and hours-of-operation.

III. Yelp’s Influence on Parents’ Child Care Search and Preferences

In order for competition to improve quality in the market for child care, consumers must be able to identify quality differences between providers. As noted in section I, the presence of information asymmetries—in which providers know more about their level of quality than do consumers—may lead consumers to make suboptimal decisions. High-quality programs are therefore not rewarded for offering such services, and low-quality programs do not have an incentive to improve (Akerlof, 1970). In this scenario, the search friction relates to consumers’ inability to obtain complete information about their child care options and/or to the high cost associated with accessing and understanding the available information.

For the purposes of this study, it is useful to think about Yelp’s collection of consumer reviews as a crowd-sourced information intervention in the child care market. By making reviews publicly available, Yelp can reduce search frictions by lowering the cost of accessing information. In addition, Yelp’s practice of revealing an average star-rating for each business further alleviates search costs by reducing the cognitive demands on consumers. One might infer, then, that the introduction of Yelp allows consumers to better distinguish between low- and high-quality child care programs, which may increase the demand for high-quality care and increase the price of utilizing it.² Furthermore, Yelp creates a powerful “feedback loop” in which providers learn from consumers what is liked and disliked about the service. Such feedback allows providers to quickly identify problems and make the necessary adjustments. Thus a Yelp-style information intervention may generate market-wide quality improvements, in part by encouraging low-quality providers to improve and by incorporating consumer feedback into their service offerings.

At issue, however, is whether child care providers and consumers are likely to respond to such information. Evidence from the early care and education (ECE) market is limited and indirect, but it suggests that consumers are increasingly comfortable with using online and peer-to-peer generated sources of information to share and shape preferences. It is noteworthy that the number of Yelp child care reviews has exploded in recent years. In our sample of 40 cities, the number of reviews per year increased from 43 in 2005 to approximately 15,000 by 2017. Furthermore, parents seem aware of and willing to use state-administered QRIS—which rate program quality and disseminate information to consumers via online portals—to aid the child care search. For example, parental awareness of QRIS ranges from 17 percent in Kentucky and 37 percent in Indiana to 87 percent in Oklahoma (Elicker et al., 2011; Star et al., 2012). In addition, two-thirds of parents in Indiana and Oklahoma state that the number of “stars” given to a program would influence their decision, and 50 percent of Indiana parents claim they would be willing to pay more for a higher-rated program.

² Of course, it is also possible that Yelp’s information disclosure will increase quality and *lower* prices by inducing competition between child care providers for consumers who use this newly-available information to alter their preferences.

Equally relevant are studies of hospitals (e.g., Jin & Sorenson, 2006) and restaurants (e.g., Jin & Leslie, 2003), which show that information interventions can alter consumer decisions and improve service delivery. A related literature covering the K-12 education sector indicates that the provision of simplified school quality information—in the form of mailers, booklets, or report cards—can improve families’ choices and student outcomes (e.g., Corcoran et al., 2018; Hastings & Weinstein, 2004). Even more relevant are the papers studying the impact of Yelp on firm quality and consumer behavior. Luca (2016) shows that an increase in a restaurant’s star-rating increases its revenue, while Anderson and Magruder (2012) find that it increases customer flows.³ Finally, the ECE literature provides suggestive evidence for the effectiveness of consumer education policies. An early paper from Chipty and Witte (1998) finds that the presence of local child care resource and referral agencies is associated with lower market prices and less price dispersion but is not related to program quality. However, a recent study by Dechausay and Anzelone (2016) finds that Indiana parents on the child care subsidy waitlist are more likely to choose high-quality arrangements if they are presented with a list of such providers located close to home.

IV. Data and Methodology

Yelp Consumer Reviews

We scraped data from Yelp by searching for pages listing “Child Care and Day Care” businesses in the 40 largest U.S. cities.⁴ We began the process by extracting the unique Yelp uniform resource locator (URL) for each business. This provided records for the universe of relevant businesses that could be referenced in case the data scraping process was interrupted.⁵ For each business, we extracted all elements available on the page. Appendix Figure 1 provides an example of a typical Yelp landing page for a child

³ A related line work, focusing on the market for health care providers, finds that Yelp is the most widely used website in the U.S. for information on hospital reviews (Bardach et al., 2013), and that hospitals’ star-ratings are correlated with objective measures of hospital quality (Bardach et al., 2013; Ranard, et al., 2016).

⁴ Yelp enables individuals to read business reviews from others and to create their own reviews. To leave a review, an individual must simply create a free account, which requires a valid email address. The individual can then rate a given business on a scale of one to five stars as well as provide a text-based review. Other individuals—irrespective of whether they have an account—can access the mobile app or website for free in order to read reviews.

⁵ Although Yelp’s business ranking methodology is not public, a visual inspection of our data shows that businesses with more reviews appear higher in the ranking than those with fewer reviews. The only apparent exception is a business that has received a recent review. For more information, see: https://www.yelp-support.com/article/How-are-search-results-ordered?l=en_US.

care business. Specifically, we scraped the business name, physical address, phone number, business website, and payment types accepted (e.g., credit cards, Apple pay, and Bitcoin). We also extracted the business’s overall Yelp rating as well as the number of 5-, 4-, 3-, 2-, and 1-star ratings. Finally, we collected information on whether the business was “claimed”. Businesses can claim their Yelp page by undergoing a verification process.⁶ A claimed Yelp business signifies that it is aware of its Yelp page and that it may be actively monitoring the page’s consumer reviews. Approximately 82 percent of the businesses in our dataset are claimed establishments.

We then extracted a variety of information on all available consumer reviews. Appendix Figure 2 provides an illustration of the layout and information available in Yelp’s child care reviews.⁷ For each review, we extracted the reviewer’s name (i.e., first name and last initial), Yelp-assigned unique ID number, city and state of residence, friend count, number of previous Yelp reviews, and whether the user has a profile picture. In addition, we captured each reviewer’s rating of the child care business (on a scale of one to five stars), the complete text-based review, and the date of the review. Finally, Yelp allows other individuals to indicate whether a given review is “useful”, “funny”, or “cool”. Thus we collected information on the number of votes received in each category. The first Yelp review in the dataset was provided in March of 2005, while the most recent review was provided in August of 2017.

Table 1 provides summary statistics for our dataset. Altogether we obtained information on 48,675 unique Yelp reviews of 9,761 child care businesses. The reviews come from 46,182 individuals. Each business received five reviews, on average. There is substantial variation across the cities in the number of reviews (and businesses) included in the data, ranging from 51 (32) in El Paso to 7,037 (791) in San Jose. The average Yelp rating is 4.3, and approximately 76 percent of ratings are five-stars. Nevertheless, there is substantial cross-city variation. The average rating ranges between 3.1 (Oklahoma City) and 4.6 (Los

⁶ <https://biz.yelp.com/support/claiming>

⁷ It is important to note that Yelp uses an automated algorithm to filter and remove from its website reviews that either violate the terms of service or are deemed fake. Nevertheless, it is possible that the Yelp filter failed to remove some fake reviews or incorrectly identified and removed authentic reviews for some businesses.

Angeles), while the share of five-star ratings ranges between 44 percent (Oklahoma City) and 88 percent (Los Angeles).

Coding the Child Care Reviews

The process for coding consumers' child care reviews was iterative and involved both deductive and inductive approaches. We began deductively by using the ECERS-R as a framework for organizing the type and frequency of quality-related language used in the text-based reviews. The ECERS-R is the most widely used observational measure of preschool-age child care classroom quality (Harms et al., 1998). As such, it has been deployed extensively by developmental psychologists to document the level and trend in child care quality. The ECERS-R is organized around seven subscales, each one containing between four and 10 items, for a total of 43 items. The items, in turn, include approximately 470 indicators to score. Trained observers rate the child care environment in relation to the full set of indicators, from which domain-specific (or subscale) scores are calculated, in addition to a single, overall measure of quality. The subscale and overall ECERS-R scores are numerical, ranging from one to seven, where one is defined as "inadequate" and seven is defined as "excellent".

Table 2 provides information on how the ECERS-R subscales were translated into words and phrases to code the Yelp reviews. Our analysis focuses on five of the seven ECERS-R subscales: space and furnishings, personal care routines, language/reasoning, activities, and interactions.⁸ We disaggregated the interactions subscale into two separate domains: general supervision of children/discipline and child-teacher interactions. We began the coding process by organizing each subscale by its constituent items and indicators. For example, the space and furnishings subscale contains eight items (e.g., indoor space; furniture for routine care, play and learning; and gross motor equipment) and over 50 individual indicators to score. For the indoor space item, for example, observers are required to determine whether there is "adequate lighting, ventilation, temperature control, and sound-absorbing materials", the "space is

⁸ We omit program structure and parents/staff. Program structure is omitted because several of its items (e.g., group time and free play) are captured sufficiently by other subscales. Parents/staff are omitted because most of the items relate to the personal and professional needs of staff, including opportunities for growth. Thus this domain is unlikely to be observed and evaluated extensively by consumers.

reasonably clean and well-maintained” and there is “good ventilation, some natural lighting through windows or skylights.” As shown in Table 2, we removed from each indicator one or more of the most salient words and phrases, so that we were left with a bank of keywords that could be used to describe each ECERS-R subscale. Note that in many cases we enhanced the word bank by including words and phrases that are closely associated with those in the actual indicator description. This keyword bank became the foundation for coding the corpus of text provided in the Yelp reviews.

Next, we utilized inductive empirical tools to further enrich the number of child care domains. Specifically, we relied on topic modelling techniques, which are unsupervised machine learning methods for finding one or more latent variables or hidden structures—referred to as “topics”—within a corpus of text. Here a topic is defined as a cluster of words that frequently co-occur within the text and that share a semantic structure. Our analysis uses the Latent Dirichlet Allocation (LDA) model to locate topics (Blei et al., 2003). LDA assumes that documents (i.e., Yelp reviews) contain a combination of topics and that topics are a distribution of words. It relies on probability distributions to determine which topics are in a given document as well as which words are in a given topic, based on word frequencies across topics and topic frequencies across the document. Using the Mallet LDA package to infer latent topics from the Yelp reviews, we experimented with different numbers of topics (i.e., 10, 20, and 30), finding that 20 generated the most meaningful ones (McCallum, 2002).⁹ We then labelled each topic based on its content.¹⁰ From this process, we identified four additional child care domains—parent interactions with management, learning and academics, referrals and recommendations, and the child care search and facility visit—and constructed a keyword bank for each one. Table 2 presents the keywords in these domains.

Finally, we created word banks for four other child care characteristics: program cost and fees, accessibility, regulations, and religious affiliation. These domains were chosen because of the strong

⁹ Mallet LDA has been widely used to analyze the content of social media in a variety of contexts (e.g., Ranard et al., 2016; Schwartz et al., 2013).

¹⁰ For example, the Mallet LDA generated a topic with the co-occurring words “place”, “rude”, “business”, “money”, “horrible”, “management”, “cares”, “turnover”, “worst”, and “fact”. These words became the basis for the child care domain “parent interactions with management”.

evidence that they are important determinants of parents’ early care and education decisions (e.g., Barbarin et al., 2006; Bassok et al., 2018b; Chaudry et al., 2011; Mamedova et al., 2013; NSECE, 2014). In addition, these domains are not captured by any of the ECERS-R subscales, nor do they not overlap with topics identified in the LDA analysis. Keywords for these domains are also shown in Table 2.

Armed with a bank of keywords and phrases in each of the 14 dimensions—seven drawn from the ECERS-R and seven additional characteristics—we then determined whether these dimensions are present in the reviews. Each dimension is represented by a vector of keywords. For example, for dimension d_1 the vector contains words such as “lights”, “furniture”, “carpet”, “sanitary”, “curtains”, “blinds”, and “playground” to capture reviews that discuss the space and furnishings subscale. A review r_i is assigned to a dimension d_j if the intersection of r_i and d_j vectors of words is not empty (i.e., $\text{vec}(r_i) \cap d_j \neq \Theta$). We created a binary indicator equal to one if a given review contains the words and phrases associated with each domain. Thus we generated 14 indicator variables. Multiple dimensions may be coded for the same review. In addition to constructing these indicators at review-level, we coded the dimensions at line-level (i.e., sentence-level) to allow for a more granular analysis of the emotional tone used by consumers when commenting on these domains.

Assessing the Affect and Emotional Tone of Reviews

We utilized two techniques to classify the linguistic and psychological properties of each Yelp review as well as the individual sentences within the review. First, we used an automated content analysis tool called Linguistic Inquiry and Word Count (LIWC), which processes and classifies textual content according to its language metrics (e.g., words per sentence), sentence structure and parts of speech (e.g., use of pronouns, verbs, articles, etc.), thinking styles (e.g., insight and certainty), and affect (e.g., positive/negative emotions, anxiety, and anger) (Pennebaker et al., 2001). As explained below, most of LIWC’s variables are expressed as a percentage of the total number of words appearing in the text. However, LIWC also produces four summary variables—analytical thinking, clout, authenticity, and emotional tone—each of which is calculated by a proprietary algorithm based on existing LIWC variables

and previous language research.¹¹ Values for these variables represent standardized scores converted to percentiles ranging from 0 to 100, where higher scores indicate a greater presence of each domain within the text. The LIWC has been used extensively in the computational linguistics and information science literatures to measure emotions expressed in unstructured, text-based data (e.g., Larsen et al., 2015; Yin et al., 2014; Taboada et al., 2011; Thelwall et al., 2011).¹²

Second, we used a supervised learning algorithm to calculate the polarity and sentiment of reviews and the individual sentences within reviews (Bontcheva et al., 2013). Supervised learning algorithms assume that the possible outputs are already known and that the data used to train the algorithm are labeled with correct answers. Thus the model can then be used to predict the labels of new unseen data. Our supervised learning algorithm uses an enhanced Naive Bayes as the classification method and mutual information as the feature selection method to calculate polarity and sentiments.¹³ Each Yelp review is assigned a value between -1 and +1, where -1 indicates a strongly negative review, +1 indicates a strongly positive review, and 0 is a neutral review. In contrast to dictionary-based sentiment algorithms, which require laborious hand-coded keywords for each domain of the dataset, our algorithm relies on a “training” dataset whose textual content is also comprised of online reviews. Our sentiment measure is strongly correlated with the business star-rating on Yelp ($r=0.54$) as well as with various LIWC measures of positive affect, including tone ($r=0.60$) and positive emotion ($r=0.32$).

¹¹ The analytical thinking variable captures the degree to which people use words that suggest formal, logical, and hierarchical thinking patterns. Clout refers to the relative social status, confidence, or leadership ability that people display through their writing or talking. Authenticity captures the degree to which people reveal themselves in an authentic or honest way, or in a more personal, humble, and vulnerable manner, through their written communication. Finally, the emotional tone variable commingles positive and negative emotions expressed in the text. The algorithm is constructed so that higher values indicate more positive tones; numbers below the 50th percentile imply increasingly negative emotional tones.

¹² The LIWC tool processes text in three steps. First, it opens and reads the input file(s) provided by the user. Second, it processes every word in the text, comparing each to a dictionary file comprising nearly 6,400 words. Each dictionary word is assigned to one or more of the grammatical and psychological categories described in the text. Finally, LIWC calculates the percentage of total words in the text corpus that matches each of the dictionary categories. To take a simple example, if we have the sentence “He hates the broccoli.” as an input to LIWC, the tool first reads “He” and compares it against the dictionary. Given that the word “He” is coded as pronoun in the dictionary, the count for this category would be incremented by one. Next, it processes “hates”, finding that it belongs to four categories: negative emotion, anger, verbs, and overall affect. Therefore, counts in these categories would be incremented by one. Once the tool processes all input text, it calculates the share of words in each category.

¹³ Naive Bayes classifier utilizes a probabilistic approach, with assumption that features are independent, to determine the labels of the unseen dataset (Narayanan et al., 2013). The algorithm applies feature selection (using mutual information) as a pre-processing step to enhance its accuracy and lower the high dimensional feature space. Feature selection is defined as the process of selecting subset of relevant features (keywords in our case) to be used in the Naive Bayes classification model.

Statistical Analysis

A key aim of the paper is to examine whether, and in what ways, consumers in lower- and higher-income markets differ in their ratings of child care providers, their preferences for various program characteristics, and the affective/emotional tones used to describe these features. We provide both descriptive and regression-based evidence on these issues. The regression model is stated as follows:

$$[1] \quad Y_{ijct} = \beta_0 + \beta_1 \ln(\text{income}_{ct}) + \mathbf{X}'\delta + \mathbf{Y}'\alpha + \lambda_1 \text{urate}_{ct} + \gamma_t + \rho_c + \varepsilon_{ijct}, \text{ where}$$

Y is some outcome pertaining to review i of firm j located in county c at time period t . The variable of interest is the log of county median household income, denoted by $\ln(\text{income})$, which is observed for each county and year included in the dataset.¹⁴ The model includes a vector of controls for user and review characteristics ($\mathbf{X}'$), including users' friend and (previous) reviews counts, whether the user has profile and review pictures posted on Yelp, whether other users rated the review as "useful", as well as the type and quality of the language used in the review (e.g., use of slang and swear words). The model also includes controls for whether the child care business is claimed and whether a given review is the first one received by the firm ($\mathbf{Y}'$). The variable *urate* denotes the unemployment rate (and a quadratic in the unemployment rate), which varies across counties and years. Finally, the model include a set of time effects (γ) and county fixed effects (ρ) to account for time-varying shocks and geographic heterogeneity that may be correlated with local household income. Standard errors in all models are clustered at the county-level.

We examine a variety outcomes in this paper. We begin by studying the impact of local household income on the business star-rating (range: one to five) using an ordered probit model, followed by the four summary language variables (range: 0 to 100) using ordinary least squares regression (OLS) as well as five affective/emotional tone outcomes (range: 0 to 1) using a fractional response model. The summary language outcomes include the measures of analytical thinking, clout, authenticity, and emotional tone.

¹⁴ We merged to the Yelp dataset county-by-year median household income over the period 2005 to 2017. This was done by first scraping the physical address of each child care business from Yelp, and then running the address listing through the TAMU Geocoding Services platform. Doing so provided a variety of geographic identifiers, including the county FIPS code for each business, which was used to merge the median household income variable (and other county-level information) to the main analysis dataset.

The affective process outcomes include positive and negative emotions, anxiety, anger, and sadness. We then turn our attention to studying the salience of the 14 child care program characteristics (i.e., the seven ECERS-R and non-ECERS-R domains), in which each characteristic is defined as a binary indicator equal to one if a review evaluates a given characteristic. We also examine whether the specific sentence(s) referencing each characteristic expresses a positive or negative sentiment. This outcome is a binary indicator equal to one if the sentence is coded as expressing a positive sentiment. These latter models are estimated using OLS. Some of our analyses are conducted at the review-level ($N=48,630$), while others are conducted at the line- or sentence-level (number of observations varies).

V. Results

Descriptive Results

We begin this section by presenting descriptive results for Yelp users' star-rating of child care businesses, the classification of consumers' child care reviews, and the affective/emotional tones expressed in the reviews. We present these data for the full sample of Yelp reviews and for reviews of businesses located in counties at the bottom and top deciles of median (county) household income. This latter analysis begins to shed light on whether consumers in poorer and wealthier markets articulate different preferences for various child care program characteristics. We pursue this question more rigorously in the next section by estimating regression models of each program characteristic on county-level household income.

Table 3 presents descriptive statistics for consumers' star-rating of child care businesses (Panel A), the four summative language variables (Panel B), the five affective process outcomes (Panel C), and the measure of positive sentiments (Panel D). The average Yelp rating is 4.3, with 13 percent of reviewers giving a one-star rating and 76 percent giving five stars. It is clear from the second and third columns that consumers in poorer markets rate their program less highly than their counterparts in wealthier markets. For example, 65 percent of consumers in the bottom income decile provide a five-star rating, while 85 percent of those in the top decile do so. Interestingly, high- and low-income consumers are about equally likely to be analytical in their reviews—that is, to use formal, logical language—but those in high-income

markets express more clout—or language characterized by confidence or leadership ability—while those in low-income markets express more authenticity—or language characterized by honesty, vulnerability, and humbleness. In addition, consumers overall use fairly positive language in their reviews: the average emotional tone score is 85 (out of 100; Panel B), and 60 percent of the individual sentences express a positive sentiment (Panel D). Nevertheless, local household income seems correlated with the use of positive emotional tones. Panels B and D show that the reviews written by high-income consumers not only score higher on the measure of emotional tone (89 versus 80), but that the individual sentences written by these consumers are more likely to express positive sentiments (63 percent versus 57 percent).

Table 4 displays the percentage of consumer reviews that comment on at least one ECERS-R subscale (Panel A) as well as a variety of other program characteristics (Panel B). Looking first at the data for the full sample, the ECERS-R domains discussed most frequently are teacher-child interactions (87 percent), personal care routines (66 percent), and activities (59 percent). The domains space and furnishings (34 percent) and language/reasoning (33 percent) are less often evaluated in the reviews. The second and third columns reveal important differences in the comments provided by poorer and wealthier consumers. It is consistently the case that consumers in wealthy markets are more likely to discuss at least one of the ECERS-R quality domains. For example, 36 percent of consumers in the top income decile mention language/reasoning in the review, while 64 percent mention activities. The comparable figures for those in the bottom income decile are 29 percent and 55 percent, respectively. In addition, higher-income consumers are more likely to evaluate the nature of their child’s interactions with teachers (91 percent versus 85 percent).

Turning to the additional program domains, displayed in Panel B, it appears that child care costs/fees (13 percent), accessibility (13 percent), and religious affiliation (two percent) garner comparatively little attention in the reviews, while learning and academics (73 percent) receives considerable attention. It is also fairly common for consumers to discuss the nature of their interactions with program managers (54 percent), their experiences during the search/facility visit process (42 percent),

and to express a recommendation (both positive and negative) about the program (45 percent). In addition, as with the ECERS-R domains, key differences emerge over the distribution of local household income. Whereas consumers in wealthy markets are more likely to evaluate the learning and academic environment (74 percent versus 70 percent) and the search/facility visitation process (45 percent versus 42 percent), those in poorer markets are more likely to discuss the program's cost and fees (16 percent versus 10 percent) and accessibility (15 percent versus 11 percent). However, the largest income difference pertains to parent interactions with management, which is included in 61 percent of reviews in the bottom decile but only 42 percent of reviews in the top decile.

The results presented in Table 5 come from the sentiment analysis of consumers' Yelp reviews. In particular, we first use the sentence- (or line-) level dataset to code each sentence as including one of the 14 domains listed in the table, and then code whether each of the domain-specific sentences expresses a positive sentiment. Thus the figures shown in Table 5 display the percentage of sentences in each child care domain that reveal a positive sentiment. Recall from our previous analysis that consumers provide overwhelmingly positive *overall* ratings of their child care provider, with 76 percent of reviews containing a five-star rating. The sentiment analysis of *specific* program domains reveals a similarly positive consumer assessment. As for the ECERS-R domains (Panel A), 82 percent of the comments on space and furnishings are positive, followed by personal care routines (79 percent), teacher-child interactions (71 percent), language/reasoning (68 percent), and activities (64 percent). The additional program features (Panel B) reveal similarly positive reactions. In fact, all seven characteristics—even the program's cost and fees—are discussed using positive language in at least 60 percent of relevant sentences. Importantly, 72 percent of the evaluations of learning and academics are positive.

However, Table 5 reveals large differences across poorer and wealthier markets in the extent to which the domain-specific reviews contain positive language. Specifically, for each of the 14 child care characteristics, consumers in wealthy markets are more likely to express positive sentiments than their counterparts in poorer markets. Such positive expressions apply to the quality of the care and learning

environment, including for example in the space and furnishings (84 percent versus 77 percent), language/reasoning (71 percent versus 64 percent), and learning/academic (73 percent versus 69 percent) domains. These differences are also present in the practical features of child care, including parent interactions with management (67 percent versus 62 percent), cost and fees (67 percent versus 65 percent), and accessibility (65 percent versus 61 percent).

Multivariate Results

We now turn to the multivariate analyses, in which we examine the impact of local household income on business' Yelp star-rating (Table 6), the emotional tones expressed in the consumer reviews (Table 6), and the child care program characteristics referenced in the reviews (Table 7).

The regression results in Table 6 are organized into four sections, with Panel A showing the results for business' Yelp rating, Panel B showing the four summative language measures, Panel C presenting the five affective process variables, and Panel D presenting the measure of positive sentiments. Results in Panels A through C are based on review-level regressions, while that in Panel D comes from a sentence-level analysis. Consistent with the descriptive results presented earlier, the estimate in Panel A reveals that a business's star-rating is strongly increasing in (the log of) county household income: consumers located in higher income communities rate their child care provider more positively. Results from the summative language variables are also consistent with the descriptive evidence: high- and low-income consumers are equally likely to be analytical in their reviews; those in higher-income markets express more clout; and those in lower-income markets express more authenticity. In addition, local household income is positively related to the use of language conveying positive affect and emotional tones. Indeed, the estimates in Panel C suggest that higher-income consumers utilize more words reflecting positive affective states; they utilize fewer words reflecting negative affective states; and they are substantially less likely to use words consistent with anger. Still more evidence that income is related to positive affect can be seen in Panel D, which shows that sentences written by consumers in higher-income communities are more likely to contain positive sentiments than the sentences written by their lower-income counterparts.

Table 7 examines the impact of local household income on two sets of outcomes: whether a given child care characteristic is referenced in consumers' Yelp reviews and whether each child care domain is evaluated using language expressing positive sentiments. Once again the child care characteristics include the ECERS-R domains (Panel A) and the auxiliary program features (Panel B). The analysis of child care domains relies on the review-level dataset, while sentiment analysis uses the sentence-level dataset.

Consistent with the descriptive results presented earlier, the first set of regression estimates reveal that lower- and higher-income consumers discuss different features of their child care provider in the Yelp reviews. Those in higher-income markets are substantially more likely to reference the quality- and education-related features of the care environment, including its space and furnishings, language/reasoning, and learning and academics. In addition, such consumers are more likely to discuss the regulatable dimensions of the child care provider (e.g., child-staff ratios and teacher education requirements). Conversely, consumers in lower-income markets are more likely to comment on the practical aspects of the child care program, including its costs and accessibility. It is also noteworthy that lower- and higher-income consumers comment on different kinds of interactions with program staff: the former is more likely discuss their interactions with managers and administrators, while the latter is more likely to focus on their child's interactions with teachers.

The second set of regressions in Table 7 provide consistent evidence that higher-income consumers discuss their child care experiences using more positive language than their lower-income counterparts. In fact, the coefficient on county household income for all 14 child care characteristics is positively signed, and is statistically significant for seven of them. Importantly, positive sentiments are more likely to be expressed by higher-income consumers for several of the quality-related domains (e.g., space and furnishings, language/reasoning, and learning and academics) as well as their interactions with both the management and teachers. In addition, the evidence suggests that such consumers have more positive experiences throughout the child care search and facility visitation process.

Yelp Users' Ratings of Review Usefulness

In this section, we examine the extent to which Yelp reviews that evaluate certain child care characteristics are rated to be “useful” by other users, and whether the number of useful votes varies by local income level.¹⁵ Such an analysis is important for a few reasons. First, all else equal Yelp elevates in its display algorithm consumer reviews that receive more useful votes by its user base. Therefore, it is important to understand whether the content of reviews is related to the number of useful votes it receives. Second, the analysis to this point has focused on the attributes of child care that are valued by *current* consumers. Indeed, by design Yelp enables current (or perhaps former) consumers to share their experiences with a child care provider. Conversely, one may reasonably assume that those voting on the usefulness of a review are more likely to be *searching* for child care. Therefore, studying the number of useful votes received by reviews in each program domain will allow for a rough comparison of preferences between *potential* and *current* child care consumers.

The data show that Yelp users are fairly active in providing useful votes. Indeed, approximately 65 percent of the child care reviews in our dataset receive at least one useful vote, with the average review receiving 4.2 votes (conditional on receiving one).¹⁶ Table 8 presents the mean number of useful votes received by reviews commenting on each of the ECERS-R and auxiliary child care characteristics, both for the full sample and for the bottom and top income deciles. It appears that, for the full sample, reviews evaluating general provider-child interactions (4.7) and program costs and fees (4.2) receive the largest number of useful votes, while those referencing personal care routines (2.9), child-teacher interactions (2.8), and learning and academics (2.8) receive the fewest, although the differences are not large. However, meaningful differences emerge once again across local income levels, with reviews of programs in the bottom decile of income consistently generating more useful votes than those in the top decile. For several

¹⁵ Yelp also allows users to vote on whether a review is “funny” and “cool”. We ultimately decided to omit these from the analysis, given the difficulties involved in interpreting their meaning as well as their relative lack of attention received by users. Indeed, on average, each review received about 0.33 funny votes and 0.37 cool votes.

¹⁶ The mean number of useful votes per review is 2.7 when reviews receiving zero votes are included in the calculation.

child care domains—particularly general provider-child interactions, program costs and fees, religious affiliation, and facility visitation—the differences are relatively large. This pattern is confirmed by estimating regressions of the number of review-level useful votes on (the log of) county household income, as shown in Table 9. Indeed, all of the coefficients on household income are negatively signed, while many are large in magnitude and statistically significant.

VI. Discussion

This study uses data on Yelp consumer reviews of child care programs in 40 U.S. cities to shed new light on parental satisfaction with and assessments of their child’s arrangement. Our analysis relies not only on data from consumers’ overall star-rating of child care businesses, but also on the information and emotional tones conveyed in the text-based reviews. A key aim of the paper is to examine whether consumer assessments of child care vary across poorer and wealthier markets. This section distills the key findings of our work, and discusses how they intersect with previous research on child care preferences.

The first important result is that consumers overall rate their child care program very highly. However, we also find substantial variation in Yelp ratings over the distribution of local household income. Still, the data show that consumers in low-income markets rate their child care program quite favorably. Insofar as Yelp’s star-rating system reflects consumers’ underlying satisfaction with their provider, our results are consistent with previous survey-based studies, all of which show that parents—even low-income ones—overwhelmingly approve of their child’s arrangement (e.g., National Public Radio, 2016; Bassok et al., 2018a; Raikes et al., 2012). It is also possible that a business’s Yelp rating is a proxy for product quality, which has either been assumed in some previous studies using Yelp data (e.g., Anderson & Magruder, 2012; Kuang, 2017; Luca & Luca, 2017) or has been tested directly in other markets (e.g., Bardach et al., 2013; Ranard, et al., 2016). To the extent that child care business ratings provide accurate signals of true quality, our results suggest that program quality may be substantially higher in wealthy communities than in poorer ones. Such a conclusion would also be consistent with findings reported in previous studies (e.g., NICHD ECCRN, 2005).

Our second key result is that consumers discuss most frequently the quality and educational dimensions of the child care provider, including the child's interactions with teachers, the learning and academic environment, and the program's personal care routines. Many of these domains have been identified by other studies as being important to parents (Bassok et al., 2018b; Barbarin et al., 2006; Chaudry et al., 2011). Less frequently discussed are topics related to the cost and availability of care, a result that contrasts with some previous work, which finds that these topics are highly salient for parents (e.g., Chaudry et al., 2011; Rose & Elicker, 2008).

In addition, this study reveals two other salient features of child care that have been overlooked in previous studies: (i) parent experiences with the child care search process and facility visit and (ii) interactions with program managers and owners. Nevertheless, the Yelp reviews show substantial variation in how parents approach the child care search. Some parents make their decision without first visiting the facility, relying instead on phone conversations with program administrators or reviewing materials posted on a company website:

I called a few places and found one that sounded good over the phone.

We chose the school, sight un-seen, by only seeing the pictures on line and reading these reviews and visiting their website, but I am soooo glad we did!! The school is precious! It's very clean, child-friendly and well maintained.

Other parents, however, visit a large number of programs before making a decision:

After interviewing with and touring roughly 20 schools, ranging from large, corporate preschools, play-based, to Montessoris, we knew that [program name omitted] was where we wanted to send our daughter.

The facility visit was clearly important to parents, as it allowed them to ask questions, learn about the curriculum, and develop a sense of comfort with and trust in the staff. For some parents, the positive first impression created by the visit solidified their decision to select the program:

When we were searching for a daycare for our first born, we looked at every single school in the area. This was the first school that took the time to answer all our questions, give us a really thorough tour (we just dropped in) and felt like a really warm, creative space for a young tot.

I randomly dropped in for a tour with my almost 1 year old daughter yesterday afternoon and was greeted by the administrators at the front desk and [name omitted] (director) jumped out and gave

me the best welcoming greeting ever. She welcomed me into her office offered me water and made me feel very comfortable, not judged at all. She was very thorough in introducing herself and [program name omitted] mission and curriculum for my daughter's age group. She took me room to room down the hall and introduced me to the teachers and let my daughter and I sit in on the classroom.

For other parents, the experience was sufficiently negative that the program was not selected:

My husband and I visited this center while looking for childcare options for our new baby. The girl at the front desk couldn't have been more than 20. She said the director was "busy" so was unable to give us a tour. The girl offered to show us around but mentioned more than once that she really didn't know too much about giving a tour, and basically just pointed out each room and told us the age group. Based on that experience, we looked elsewhere for childcare.

In terms of consumer interactions with program managers and owners, the Yelp reviews reveal that parents expect directors, in particular, to fulfill multiple, complex roles as administrators and managers as well as pedagogues. It is clear that parents use their meetings with program directors to form a first and lasting impression of the program as a whole, and, as such, it is common for parents to assess the director's warmth and compassion, educational philosophy, and business savvy. Interestingly, many parents comment on the "customer service" aspects of their interaction with the director, while others evaluate the director's organizational abilities and responsiveness to questions:

The director really understands how to provide a quality program, select qualified, knowledgeable, and fun teachers, and give families excellent customer service.

As daycare owners in the area go, [name omitted] is definitely a cut above: she responds promptly to emails and calls or texts, follows up on questions and is very well-organized.

In addition, parents evaluate owners and managers in terms of their ability to increase program quality, as measured by a range of factors, from improving a facility's physical appearance to lowering staff turnover:

Needless to say my fears were quickly dispelled when the new owner immediately spruced up the building and grounds, eliminating thirty years of accumulated clutter so that the children have a bright, clean place in which to play and learn.

The teaching staff was always good, but with the new owners, the turnover has slowed and both of my girls always talk about moving to the next class to see [name omitted].

Finally, we encountered some reviews in which parents' describe their interaction with the director as being so positive that it became the central reason for selecting the provider:

We were going to enroll our daughter at a preschool with a much higher tuition fee but we didn't like how the facility felt like for us. We chose [program name omitted] because [name omitted] (owner) made us feel very welcome. When I did the tour with my daughter she was very attentive to her and I saw a quick connection with both.

The third important result is that lower- and higher-income consumers evaluate different child care characteristics in their Yelp reviews. The former is more likely to comment on a program's practical features, such as its pricing and accessibility, while the latter is more likely to discuss the learning and academic environment. Although this income-based split in parental assessments has been identified elsewhere (e.g., Rose & Elicker, 2008), we find that one of the largest differences occurs in a program domain that has not been studied previously: parent interactions with program managers. Indeed, our descriptive evidence reveals that 61 percent of consumers in the bottom income decile evaluate this characteristic, while 42 percent of those in the top decile do so. We also find that consumers in higher-income markets describe their child care experiences using substantially more positive language. Generally speaking, higher-income consumers are more likely to use language consistent with positive emotional tones, while lower-income consumers are more likely to express a range of negative affective states, such as anxiety and anger. More specifically, wealthier consumers provide comparatively positive assessments not only of the structural (e.g., space and furnishings) and process (e.g., teacher-child interactions) features of the child care environment, but also of the search and facility visitation process.

Although lower-income consumers express many concerns in their Yelp reviews, two appear to be particularly salient. First, many parents describe chaotic environments in which children hurt one another, teachers fail to bring order to the classroom (and were sometimes rough with children), and children are not provided with enriching activities:

I enrolled my 2 yr old daughter here and pulled her out after just two days. Teachers were not paying attention at all as kids are hurting each other. Did not comfort crying children. Director was impossible to reach and unapologetic...

My child had a bully here and she was getting hit and had her hair pulled. The staff did nothing even tho my daughter came home with marks. I came to pick her up and she was being put in time out for what the other child did.

Related concerns for many parents are high teacher turnover and inadequate supervision:

New teachers every week. Children are moved from classroom to classroom! I never know who is really watching my child. Walked into a 2 year old class and children were everywhere. I would say upwards of twenty for two adults.

The second key concern for low-income parents focuses on the organizational ability and communication style of program administrators. Lower-income parents often criticize directors and owners for being disorganized, for communicating disrespectfully, and for placing profit ahead of safety and quality:

I am giving [program name omitted] 2 stars based on organization and communication. I kept my 8 mo. daughter in mainly because I really loved one of the teachers in the class. My main problem was with the management so this rating is based on that.

This school is ran very poorly. Teachers are good but administration is terrible. If there's any kind of major issue such as bullying or harassment [name omitted] will do nothing about it. In fact she will get defensive and lose her temper at parents.

One drawback of many studies within the child care choice literature is that parent surveys of program preferences are typically administered after the family has selected the child's arrangement. As a result, parent responses likely reflect some combination of their true preferences as well as the financial and supply constraints they face in the market. To some extent, our analysis using Yelp reviews suffers from the same critique: consumers lend their insights on what program features matter most only after the child has enrolled in the program. This may explain the somewhat surprising result that low percentages of consumers discuss program costs and accessibility in the reviews. These features are likely to be known beforehand, and they powerfully shape the choice set faced by consumers. Thus characteristics like cost and accessibility are less salient to parents after the child enrolls in the program. However, Yelp allows users to rate whether a given review is "useful". Under the assumption that these ratings are driven by parents searching for—rather than currently using—child care, variation in useful votes received across each of the 14 domains may reasonably approximate parental preferences *prior* to consumption.

Two results from our analysis of Yelp's useful votes are noteworthy. First, reviews assessing general supervision and discipline, program costs, and accessibility received the largest number of useful votes. This stands in contrast to the share of reviews commenting on these domains. Indeed, these were not among the most heavily discussed topics in the reviews. Such a disconnect supports the notion that

cost and accessibility are highly salient program features during the child care search, but become less so over time as characteristics like facility cleanliness, activities and curriculum, and child-teacher interactions increase in salience after the child is exposed to the program. Second, we find that reviews posted in lower-income markets receive more useful votes in every domain than those in higher-income markets. Again, this conflicts with the result that consumers in lower-income markets are less likely to discuss most program features in their reviews. One explanation for the disconnect is that low-income families face more severe information problems: limited social networks or awareness of QRIS (or resource and referral organizations) may preclude learning about local child care options. Thus Yelp may be an important information resource for disadvantaged families. However, once these families make their child care decision, they develop different valuations of the relative importance of various program features.

VII. Conclusions

Our results indicate that the ECERS-R is reasonably effective at capturing the contents of parents' own evaluation of their child care program. On the one hand, a majority of consumers assess in their Yelp reviews the ECERS-R subscales for child-teacher interactions, personal care routines, and activities. However, relatively few consumers evaluate the space and furnishings and language/reasoning subscales. Furthermore, the ECERS-R does not capture at all a number of child care domains that are highly salient to consumers, including the learning and academic environment, the nature of parents' interactions with management, and experiences with the search and facility visit process. Although the ECERS-R was not designed to measure some of these latter domains, it is nonetheless clear that parents have preferences for a complex mix of practical and quality-related characteristics that extend beyond those reflected in standard measures of program quality.

We also find that consumers in poorer communities give lower ratings to their provider; they are less likely to evaluate many of the most important quality-related features of their program; and when they do discuss these features, they are described in more negative terms. Such results indicate that low-income families may be using lower-quality care and are less satisfied with this care than their high-income

counterparts. Thus there may be a role for public policy via consumer education and subsidies to assist disadvantaged families. One approach, taken in most states, is to adopt a QRIS, which provides a structure for assessing and publicly reporting program quality, so that parents can make informed decisions.

For at least two reasons, however, consumer education policies such as QRIS are not a panacea. First, as others note, QRIS may be effective at changing preferences around high-quality child care, but these systems are not designed to alleviate gaps in supply, nor will they increase parents' ability to pay for higher-quality care (Bassok et al., 2018b). Therefore, policies like QRIS may not lead to observationally different choices if parents remain geographically and financially precluded from accessing better programs. Thus a system that couples consumer education with financial assistance may be necessary to lead to better choices. Second, results from this study suggest that QRIS, at least as it is structured in many states, does not capture some features of child care that are important to parents, including the business and customer service aspects of the search as well as the nature of parents' interactions with management. Lower-income families, in particular, often report negative experiences with both characteristics. Thus some attempt could be made to incorporate these program features into the QRIS ratings.

Our study makes several contributions to the literature. The use of consumer review data from Yelp provides an opportunity to examine parent preferences on a national scale, across a demographically and economically diverse set of communities, and for a large and diverse set of child care programs. Indeed, most prior studies are constrained to a small number of locales (e.g., Bassok et al., 2018a,b; Mocan, 2007), to an economically homogenous set of families (e.g., Bassok et al., 2018b; Raikes et al., 2012), or to a single program-type (e.g., Barbarin et al., 2006). The diversity of consumers in the Yelp data is particularly important, given the differences we document in child care preferences across lower- and higher-income markets. The richness of Yelp's consumer data is another advantage. Yelp provides not only an overall measure of consumer satisfaction—by way of its star rating—but it allows users to offer narrative-based program reviews. Such textual data offer some advantages over the standard data collection approach in the literature (i.e., parent surveys) because it allows for open-ended descriptions of a program's most salient

features without the constraints of researchers' *a priori* beliefs about what defines "high-quality" child care driving the survey's construction. Furthermore, the sheer volume of available data allowed us to examine the salience of a broad range of program characteristics. This includes features captured by environmental rating scales of classroom quality (e.g., ECERS-R) as well as a number of non-quality features, many of which are shown here to be important to parents' evaluation of child care. A final contribution of the paper is its analysis of parents' affect when assessing various program characteristics. We provide consistent evidence that lower-income consumers are less satisfied, more anxious, and angrier about many aspects of their child care experience. Detection of these negative experiences, which may reflect the low quality of care used by these families, is made possible by the open-ended nature of Yelp's consumer reviews as well as the use of powerful text analysis techniques.

References

- Akerlof, G.A. (1970). The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 84, 488-500.
- Anderson, M. & Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *Economic Journal*, 122, 957-989.
- Anderson, S., Ramsburg, D. M., & Scott, J. (2005). *Illinois study of license-exempt child care: Final report*.
- Auger, A., Farkas, G., Burchinal, M., Duncan, G., & Vandell, D. (2014). Preschool center care quality effects on academic achievement: An instrumental variables analysis. *Developmental Psychology*, 50, 2559-2571.
- Barbarin, O. A., McCandies, T., Early, D., Clifford, R. M., Bryant, D., Burchinal, M., Howes, C., ...Pianta, R. (2006). Quality of prekindergarten: What families are looking for in public sponsored programs. *Early Education and Development*, 17, 619-642.
- Bardach NS, Asteria-Peñaloza R, Boscardin WJ, Dudley RA. (2013). The relationship between commercial website ratings and traditional hospital performance measures in the USA. *BMJ Qual Saf*, 22, 194-202.
- Bassok, D., Fitzpatrick, M., Loeb, S. & Paglayan, A. (2013). The early childhood care and education workforce from 1990 through 2010: Changing dynamics and persistent concerns. *Education Finance and Policy*, 8, 581-601.
- Bassok, D., Markowitz, A., Player, D., & Zagardo, M. (2018a). Are parents' ratings and satisfaction with preschools related to program features? *AERA Open*, 4, 1-17.
- Bassok, D., Magouirk, P., Markowitz, A., Player, D. (2018b). Are there differences in parents' preferences and search processes across preschool types? Evidence from Louisiana. *Early Childhood Research Quarterly*, 44, 43-54.
- Blau, D. (2001). *The Child Care Problem: An Economic Analysis*. Russell Sage Foundation.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Bontcheva, K., Derczynski, L., Funk, A., Greenwood, M., Maynard, D., & Aswani, N. (2013). TwitIE: An open-source information extraction pipeline for microblog text. In *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013* (pp. 83-90).
- Boyd-Swan, C. & Herbst, C.M. (2018). The Demand for Teacher Characteristics in the Market for Child Care: Evidence from a Field Experiment. *Journal of Public Economics*, 159, 183-202.
- Chaudry, A., Pedroza, J., Sandstrom, H., Danziger, A., Grosz, M., Scott, M.M., & Ting, S. (2011). *Child care choices of low income working families* (pp. 1-212). Washington, DC: Urban Institute.
- Chipty, T. & Witte, A.D. (1998). Effects of information provision in a vertically differentiated market. NBER Working Paper No. 6493. Cambridge, MA: National Bureau of Economic Research.

Corcoran, S., Jennings, J., Cohodes, S., & Sattin-Bajaj, C. (2018). Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions. NBER Working Paper No. 24471. Cambridge, MA: National Bureau of Economic Research.

Cryer, C. & Burchinal, M. (1997). Parents as child care consumers. *Early Childhood Research Quarterly*, 12, 35-58.

Cryer, C., Tietze, W., & Wessels, H. (2002). Parents' perceptions of their children's child care. A cross-national comparison. *Early Childhood Research Quarterly*, 17, 259-277.

Dechausay, N. & Anzelone, C. (2016). Cutting through Complexity: Using Behavioral Science to Improve Indiana's Child Care Subsidy Program. OPRE Report 2016-03. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Early, D. M., & Burchinal, M. R. (2001). Early childhood care: Relations with family characteristics and preferred care characteristics. *Early Childhood Research Quarterly*, 475-497.

Elicker, J., Langill, C., & colleagues. (2011). Evaluation of Paths to QUALITY, Indiana's Child Care Quality Rating and Improvement System: Final Report. West Lafayette, IN: Purdue University.

Forry, N. D., Tout, K., Rothenberg, L., Sandstrom, H., Vesely, C. (2013a). Child Care Decision-Making Literature Review. OPRE Brief 2013-45. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Forry, N., Simkin, S., Wheeler, E., & Bock, A. (2013b) 'You know how it makes you feel': Low-income parents' childcare priorities and definitions of ideal high-quality childcare. *Journal of Children and Poverty*, 19, 107-126.

Forry, N. D., Isner, T. K., Daneri, P., & Tout, K. (2014). Child care decision-making: Understanding priorities and processes used by subsidized low-income families in Minnesota. *Early Education and Development*, 25, 995-1015.

Gathman, C & Sass, B. (2017). Taxing childcare: Effects on childcare choices, family labor, supply and children. Forthcoming at *Journal of Labor Economics*.

Gordon, R. A., & Hognas, R. S. (2006). The best laid plans: Expectations, preferences, and stability of child-care arrangements. *Journal of Marriage & the Family*, 68, 373-393.

Harms, T., Clifford, R. M., & Cryer, D. (1998). Early Childhood Environment Rating Scale: Revised Edition. New York, NY: Teachers College Press.

Hastings, J. S., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *Quarterly Journal of Economics*, 123, 1373-1414.

Havnes, T. & Mogstad, M. (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy*, 3, 97-129.

- Herbst, C.M. (2013). The impact of non-parental child care on child development: Evidence from the summer participation “dip”. *Journal of Public Economics*, 105, 86-105.
- Herbst, C.M. (2017). Universal child care, maternal employment, and children’s long-run outcomes: Evidence from the U.S. Lanham Act of 1940. *Journal of Labor Economics*, 35, 519-564.
- Herbst, C.M. (2018). The rising cost of child care in the United States: A reassessment of the evidence. *Economics of Education Review*, 64, 13-30.
- Jin, G. & Leslie, P. (2003). The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards. *Quarterly Journal of Economics*, 118, 409-451.
- Jin, G. & Sorensen, A. (2006). Information and consumer choice: The value of publicized health plan ratings. *Journal of Health Economics*, 25, 248-275.
- Johansen, A. S., Leibowitz, A., & Waite, L. J. (1996). The importance of child-care characteristics to choice of care. *Journal of Marriage and Family*, 58, 759-772.
- Keys, T. D., Farkas, G., Burchinal, M. R., Duncan, G. J., Vandell, D. L., Li, W., Ruzek, E. A. & Howes, C. (2013). Preschool Center Quality and School Readiness: Quality Effects and Variation by Demographic and Child Characteristics. *Child Development*, 84, 1171-1190.
- Kuang, C. (2017). Does quality matter in local consumption amenities? An empirical investigation iwht Yelp. *Journal of Urban Economics*, 100, 1-18.
- Larsen, M. E., Boonstra, T. W., Batterham, P. J., O’Dea, B., Paris, C., & Christensen, H. (2015). We feel: mapping emotion on Twitter. *IEEE Journal of Biomedical and Health Informatics*, 19, 1246-1252.
- Layzer, J. I., Goodson, B. D., & Brown-Lyons, M. (2007). *National Study of Care for Low-Income Families: Care in the home: A description of family child care and the experiences of the families and children that use it: Final Report*.
- Luca, M. (2016). Reviews, reputation, and revenue: The case of Yelp.com. Working Paper No. 12-016. Cambridge, MA: Harvard Business School.
- Luca, D.L. & Luca, M. (2017). Survival of the fittest: The impact of the minimum wage on firm exit. Working Paper No. 17-088. Cambridge, MA: Harvard Business School.
- McCallum, A. K. (2002). MALLET: A Machine Learning for Language Toolkit. <http://mallet.cs.umass.edu>.
- Mamedova, S. & Redford, J. (2013). Early Childhood Program Participation, From the National Household Education Surveys Program of 2012 (NCES 2013-029.REV), National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Mocan, N. (2007). Can consumers detect lemons? An empirical investigation of information asymmetry in the market for child care. *Journal of Population Economics*, 20, 507-531.

National Research Council & Institute of Medicine (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC: National Academies Press.

Narayanan, V., Arora, I., & Bhatia, A. (2013). Fast and accurate sentiment classification using an enhanced Naive Bayes model. In *International Conference on Intelligent Data Engineering and Automated Learning* (pp. 194-201). Springer, Berlin, Heidelberg.

National Public Radio. (2016). Child care and health in America. Available at: <https://www.npr.org/sections/health-shots/2017/01/03/506448993/child-care-scarcity-has-very-real-consequences-for-working-families>.

National Survey of Early Care and Education Project Team (NSECE). (2014). Household search for and perceptions of early care and education: Initial findings from the National Survey of Early Care and Education (NSECE) (OPRE Report No. 2014-55a). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

NICHD ECCRN. (1997). Familial factors associated with the characteristics of nonmaternal care for infants. *Journal of Marriage and Family*, 59, 389-408.

NICHD ECCRN. (2005). Characteristics and quality of child care for toddlers and preschoolers. Chapter 6. *Child Care and Child Development: Results from the NICHD Study of Early Child Care and Youth Development*. New York, NY: Guilford.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*, 71.

Peyton, V., Jacobs, A. C., O'Brien, M., & Roy, C. (2001). Reasons for choosing child care: Associations with family factors, quality and satisfaction. *Early Childhood Research Quarterly*, 16, 191-208.

Pungello, E. P., & Kurtz-Costes, B. (2000). Working women's selection of care for their infants: A prospective study. *Family Relations*, 49, 245-255.

Raikes, H., Torquati, J., Wang, C., & Shjegstad, B. (2012). Parent Experiences with State Child Care Subsidy Systems and Their Perceptions of Choice and Quality in Care Selected. *Early Education and Development*, 23, 558-582.

Ranard BL, Werner RM, Antanavicius T, Schwartz HA, Smith RJ, Meisel ZF, Asch DA, Ungar LH, Merchant RM. (2016). Yelp Reviews Of Hospital Care Can Supplement And Inform Traditional Surveys Of The Patient Experience Of Care. *Health Affairs*, 35, 697-705.

Rose, K. K., & Elicker, J. (2008). Parental decision making about child care. *Journal of Family Issues*, 29, 1161-1184.

Rossin-Slater, M. & Wust, M. (2016). What is the added value of preschool? Long-term impacts and interactions with a health intervention. IZA Discussion Paper No. 10254.

Sandstrom, H., & Chaudry, A. (2012). 'You have to choose your childcare to fit your work': Childcare decision-making among low-income working families. *Journal of Children and Poverty*, 18, 89-119.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... Ungar, L. H. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLoS ONE*, 8, e73791.

Shlay, A. B., Tran, H., Weinraub, M., & Harmon, M. (2005). Teasing apart the child care conundrum: A factorial survey analysis of perceptions of child care quality, fair market price and willingness to pay by low-income, African American parents. *Early Childhood Research Quarterly*, 20, 393-416.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37, 267-307.

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the Association for Information Science and Technology*, 62, 406-418.

Yin, D., Bond, S., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly*, 2, 539-560.

Table 1: Summary Statistics on the Yelp Child Care Reviews, by City

City	Number of Businesses	Number of Reviews	Number of Reviewers	Average Yelp Rating	Percent 5-Stars
Albuquerque	43	90	84	3.68	0.577
Atlanta	168	536	519	4.15	0.725
Austin	224	1,801	1,676	4.33	0.754
Baltimore	76	138	137	3.60	0.572
Boston	267	761	739	4.10	0.712
Charlotte	109	221	207	3.65	0.570
Chicago	525	2,641	2,481	4.24	0.741
Colorado Springs	34	71	69	3.35	0.521
Columbus	100	186	179	3.83	0.650
Dallas	344	1,083	1,018	3.82	0.627
Denver	290	928	874	4.00	0.672
Detroit	128	252	244	3.87	0.654
El Paso	32	51	45	3.82	0.607
Fort Worth	188	609	583	3.71	0.597
Fresno	45	141	132	4.30	0.751
Houston	429	1,361	1,274	3.71	0.593
Indianapolis	78	144	143	3.99	0.659
Jacksonville	87	167	160	3.71	0.568
Kansas City	113	189	178	3.75	0.576
Las Vegas	170	1,253	1,163	3.84	0.643
Long Beach	356	2,063	1,987	4.25	0.763
Los Angeles	786	6,199	6,003	4.64	0.879
Louisville	40	59	58	3.33	0.525
Memphis	37	71	67	3.42	0.478
Mesa	59	121	119	3.80	0.619
Milwaukee	64	92	86	3.61	0.608
Nashville	86	182	171	3.54	0.549
New York	621	2,489	2,388	4.34	0.764
Oklahoma City	46	80	75	3.13	0.437
Philadelphia	253	597	566	3.69	0.603
Phoenix	359	1,846	1,736	4.01	0.701
Portland	299	1,113	1,051	4.19	0.734
Sacramento	383	2,154	2,021	4.33	0.772
San Antonio	194	459	423	3.53	0.542
San Diego	524	3,629	3,431	4.44	0.815
San Francisco	650	4,817	4,558	4.56	0.843
San Jose	791	7,037	6,631	4.55	0.835
Seattle	441	2,105	1,990	4.05	0.675
Tucson	57	146	139	3.45	0.561
Washington, D.C.	265	793	777	3.93	0.646
Total	9,761	48,675	46,182	4.29	0.762

Table 2: Framework for Classifying Yelp Child Care Reviews

Dimension	Sample ECERS-R Items	Sample Word Bank	Sample Line of Text
Panel A: ECERS-R Subscales			
Space and furnishings	- Indoor space - Furnishings for care/play/learning - Space and equipment for gross motor play	facility, space, class, room, lights, furniture, carpet, curtains, playground, shade, odor, bright, clean	“You can just tell that the kids are happy, the facility is clean and bright, and everyone is so friendly!”
Personal care routines	- Greeting and departing - Meals and snacks - Health practices	greet, smiles, hugs, allergies, meal, snack, food, nutritious, germs, wash, toilet, crib, nap/rest	“When the children arrive they are always greeted with a big Buenos Dias! and a hug if the child is willing...”
Language/Reasoning	- Books and pictures - Encouraging children to communicate	books, read, social, interact	“What's more, my son was reading books and writing words at 4 years old without any intervention on my part because the Montessori method moves each child along at their own pace...”
Activities	- Fine motor - Art - Dramatic play	toys, activities, play, art, blocks, puzzles, crayons, music, sing, dance	“When I return at 5 they are always involved and engaged in creative and constructive activities such as puzzles, blocks, role play or even preposition bingo.”
Interaction	- General supervision of children - Discipline - Child-to-child interactions - Interactions with teachers	supervise, discipline, yell, spank, tease, bully, caring, warm, patient, loving	“Our daughter was there for 2 days only and we had to stop sending her there after noticing how physically rough one of the teacher's assistant was with her.”
Panel B: Additional Child Care Program Features			
Parent interactions with management	N/A	owner, business, management, director, service, rude	“The owner is horrible at customer service!!! We've ask the front desk to call us over 6 times to discuss changing our child's days of the week”
Learning and academics	N/A	curriculum, learn, teach development, skills, preschool, education	“The teachers are sweet and caring, and the curriculum always surprises me with what they can teach a 2 year old - colors, numbers, shapes, French, music...”
Referrals and recommendations	N/A	refer/referral, recommend, decision, sending, review, stars	“Very grateful I have found a safe and happy space for my children and highly recommend!!!”
Search and facility visit	N/A	tour, visit, questions, schedule, phone	“I went and toured and was disgusted- tv's in the daycare, small dirty spaces and very rude staff”
Program cost and fees	N/A	price, cost, fees, tuition, affordable, expensive	“A few things to add about X Center are 1 the meals provided are tasty, balanced,...2

			the tuition is very reasonable and one of the most affordable in the area.”
Accessibility	N/A	accessible, convenient, hours, location, time, day, open, close, home, work	“Unlike most other daycares in the city, their hours are 7:30am to 6:30pm perfect for working parents.”
Regulation features	N/A	regulations, ratios, inspection, violation, credentials, certified, background (check), fingerprints	“Although it is a big facility, it is clean and adequately staffed with appropriate children teacher ratios.”
Religious affiliation	N/A	religion, religious, Christian, Catholic, Jewish	“It is a Christian-based preschool in a Presbyterian Church - and they teach fundamental Christian values and lessons, but by no means ostracize those who are not.”

Table 3: Yelp Business Rating and LIWC Language Variables

Variable	Full Sample	Bottom Income Decile	Top Income Decile
<i>Panel A: Yelp Business Rating</i>			
Yelp star rating (ave.)	4.29 (1.42)	3.98 (1.59)	4.56 (1.17)
Yelp rating: 1 star (%)	0.133 (0.340)	0.191 (0.394)	0.081 (0.273)
Yelp rating: 5 stars (%)	0.762 (0.426)	0.651 (0.477)	0.854 (0.353)
<i>Panel B: Summary Language Variables</i>			
Analytic thinking	51.59 (22.78)	51.65 (23.02)	53.21 (22.42)
Clout	77.44 (19.09)	74.20 (20.27)	81.28 (17.07)
Authenticity	42.78 (27.80)	46.42 (27.92)	38.86 (26.99)
Emotional tone	84.64 (25.87)	79.77 (29.58)	89.16 (21.11)
<i>Panel C: Affective Processes</i>			
Positive emotions, overall (%)	0.068 (0.038)	0.063 (0.040)	0.073 (0.036)
Negative emotions, overall (%)	0.009 (0.013)	0.010 (0.015)	0.007 (0.010)
Anxiety (%)	0.002 (0.005)	0.002 (0.005)	0.002 (0.004)
Anger (%)	0.001 (0.005)	0.002 (0.006)	0.001 (0.004)
Sadness (%)	0.003 (0.006)	0.003 (0.007)	0.003 (0.006)
<i>Panel D: Sentiment Measure</i>			
Positive sentiment, line-level (%)	0.599 (0.490)	0.566 (0.496)	0.633 (0.482)

Notes: Standard deviations are shown in parentheses.

Table 4: Classification of Consumers' Child Care Business Reviews

Share of Reviews in Each Category	Full Sample	Bottom Income Decile	Top Income Decile
<i>Panel A: ECERS-R Subscales</i>			
Space and furnishings (%)	0.341 (0.474)	0.333 (0.471)	0.382 (0.486)
Personal care routines (%)	0.645 (0.479)	0.624 (0.484)	0.697 (0.459)
Language/Reasoning (%)	0.328 (0.469)	0.285 (0.451)	0.358 (0.479)
Activities (%)	0.592 (0.491)	0.549 (0.498)	0.642 (0.480)
Interactions, supervision/discipline (%)	0.104 (0.305)	0.117 (0.321)	0.106 (0.308)
Interactions, teachers (%)	0.873 (0.334)	0.849 (0.358)	0.908 (0.289)
<i>Panel B: Additional Child Care Program Features</i>			
Parent interactions with management (%)	0.544 (0.498)	0.613 (0.487)	0.421 (0.494)
Learning and academics (%)	0.734 (0.442)	0.698 (0.459)	0.743 (0.437)
Referrals and recommendations (%)	0.451 (0.498)	0.418 (0.493)	0.496 (0.500)
Search and facility visit (%)	0.424 (0.494)	0.416 (0.493)	0.448 (0.497)
Program cost and fees (%)	0.125 (0.331)	0.164 (0.370)	0.099 (0.299)
Accessibility (%)	0.129 (0.335)	0.149 (0.356)	0.110 (0.313)
Regulation features (%)	0.221 (0.415)	0.228 (0.420)	0.206 (0.404)
Religious affiliation (%)	0.018 (0.134)	0.019 (0.136)	0.013 (0.113)

Notes: Standard deviations are shown in parentheses.

Table 5: Sentiment Analysis of Child Care Program Features

Share of Lines in Each Category Coded “Positive” Sentiment	Full Sample	Bottom Income Decile	Top Income Decile
<i>Panel A: ECERS-R Subscales</i>			
Space and furnishings (%)	0.816 (0.387)	0.766 (0.423)	0.836 (0.370)
Personal care routines (%)	0.785 (0.411)	0.752 (0.432)	0.812 (0.391)
Language/Reasoning (%)	0.684 (0.465)	0.644 (0.479)	0.705 (0.456)
Activities (%)	0.642 (0.479)	0.608 (0.488)	0.656 (0.475)
Interactions, supervision/discipline (%)	0.542 (0.498)	0.489 (0.500)	0.639 (0.481)
Interactions, teachers (%)	0.708 (0.455)	0.672 (0.470)	0.740 (0.438)
<i>Panel B: Additional Child Care Program Features</i>			
Parent interactions with management (%)	0.651 (0.477)	0.618 (0.486)	0.671 (0.470)
Learning and academics (%)	0.715 (0.452)	0.690 (0.463)	0.728 (0.445)
Referrals and recommendations (%)	0.612 (0.487)	0.573 (0.495)	0.645 (0.479)
Search and facility visit (%)	0.602 (0.489)	0.569 (0.495)	0.637 (0.481)
Program cost and fees (%)	0.660 (0.474)	0.652 (0.476)	0.670 (0.470)
Accessibility (%)	0.627 (0.484)	0.611 (0.488)	0.646 (0.479)
Regulation features (%)	0.621 (0.485)	0.581 (0.494)	0.632 (0.482)
Religious affiliation (%)	0.607 (0.489)	0.571 (0.497)	0.705 (0.459)

Notes: Standard deviations are shown in parentheses.

Table 6: The Relationship between Local Household Income and Yelp Business Rating and LIWC Language Variables

Dependent Variable	Coeff./SE	Observations
<i>Panel A: Yelp Business Rating</i>		
Yelp star rating	0.890*** (0.195)	48,630
<i>Panel B: Summary Language Variables</i>		
Analytic thinking	-0.696 (3.701)	48,630
Clout	17.397*** (2.951)	48,630
Authenticity	-9.981*** (3.675)	48,630
Emotional tone	12.062*** (3.395)	48,630
<i>Panel C: Affective Processes</i>		
Positive emotions, overall	0.239*** (0.077)	48,630
Negative emotions, overall	-0.486* (0.275)	48,630
Anxiety	-0.648 (0.451)	48,630
Anger	-1.217*** (0.417)	48,630
Sadness	0.223 (0.413)	48,630
<i>Panel D: Sentiment Measure</i>		
Sentiment, line-level	0.419*** (0.108)	438,361

Notes: Each cell presents the coefficient on the log of county median household income and its standard error (in parentheses) clustered at the county-level. Each coefficient and standard error come from a separate regression. The estimators used are ordered probit (first outcome), ordinary least squares regression (second through fifth outcomes), GLM for fractional responses (sixth through tenth outcomes), and ordered probit (eleventh outcome). All models include the full set of controls listed in Appendix Table 1, calendar quarter dummies, and county fixed effects. *, **, and *** indicate that a given coefficient is statistically significant at the 0.10, 0.05, and 0.01 level, respectively.

Table 7: The Relationship between Local Household Income and the Classification of Consumers' Child Care Business Reviews

	DV: Review Category		DV: Sentiment Measure	
	Coeff./SE	Observations	Coeff./SE	Observations
<i>Panel A: ECERS-R Subscales</i>				
Space and furnishings	0.141* (0.075)	48,630	0.680** (0.321)	22,346
Personal care routines	0.085 (0.071)	48,630	0.302* (0.182)	59,726
Language/Reasoning	0.216** (0.093)	48,630	0.523*** (0.180)	23,105
Activities	0.108* (0.061)	48,630	0.178 (0.171)	57,313
Interactions, supervision/discipline	0.065* (0.038)	48,630	0.266 (0.523)	6,522
Interactions, teachers	0.251*** (0.037)	48,630	0.487*** (0.129)	123,132
<i>Panel B: Additional Child Care Program Features</i>				
Parent interactions w/management	-0.156** (0.071)	48,630	0.342* (0.186)	46,433
Learning and academics	0.138* (0.074)	48,630	0.349** (0.152)	97,687
Referrals and recommendations	0.274*** (0.055)	48,630	0.212 (0.199)	28,322
Search and facility visit	0.185*** (0.068)	48,630	0.415* (0.224)	34,305
Program cost and fees	-0.091** (0.040)	48,630	0.484 (0.410)	7,464
Accessibility	-0.084* (0.051)	48,630	0.380 (0.322)	7,788
Regulation features	0.106** (0.051)	48,630	0.260 (0.314)	14,452
Religious affiliation	0.006 (0.025)	48,630	1.260 (1.367)	1,149

Notes: Each cell presents the coefficient on the log of county median household income and its standard error (in parentheses) clustered at the county-level. The first set of models is estimated on the review-level dataset; the second set of models is estimated on the line-level dataset. All models include the full set of controls listed in Appendix Table 1, calendar quarter dummies, and county fixed effects. *, **, and *** indicate that a given coefficient is statistically significant at the 0.10, 0.05, and 0.01 level, respectively.

Table 8: Yelp Users' Ratings of Review 'Usefulness'

Mean Number of "Useful" Votes for Reviews in Each Category	Full Sample	Bottom Income Decile	Top Income Decile
<i>Panel A: ECERS-R Subscales</i>			
Space and furnishings	3.044 (5.523)	4.813 (6.994)	1.895 (4.050)
Personal care routines	2.884 (5.191)	4.486 (6.604)	1.807 (3.773)
Language/Reasoning	3.349 (5.828)	5.069 (7.551)	2.227 (4.609)
Activities	3.014 (5.411)	4.767 (6.973)	1.975 (4.063)
Interactions, supervision/discipline	4.680 (7.604)	6.394 (9.534)	2.910 (5.392)
Interactions, teachers	2.754 (5.082)	4.281 (6.266)	1.758 (3.633)
<i>Panel B: Additional Child Care Program Features</i>			
Parent interactions with management	3.205 (5.550)	4.619 (6.907)	2.292 (4.433)
Learning and academics	2.831 (5.242)	4.459 (6.560)	1.884 (3.910)
Referrals and recommendations	3.055 (5.390)	4.736 (6.859)	1.939 (4.065)
Search and facility visit	3.358 (5.786)	5.057 (7.421)	2.046 (3.758)
Program cost and fees	4.154 (7.103)	6.132 (8.342)	2.512 (4.103)
Accessibility	3.798 (6.393)	5.432 (7.311)	2.602 (4.718)
Regulation features	3.896 (6.741)	5.529 (8.057)	2.778 (5.098)
Religious affiliation	3.449 (7.884)	5.515 (6.651)	2.145 (3.208)

Notes: Standard deviations are shown in parentheses.

Table 9: The Relationship between Local Household Income and the Number of “Useful” Votes Received in Each Category

	Coeff./SE	Observations
<i>Panel A: ECERS-R Subscales</i>		
Space and furnishings	-7.243** (2.955)	16,604
Personal care routines	-5.614** (2.491)	31,354
Language/Reasoning	-7.099** (3.053)	15,941
Activities	-5.717* (2.967)	28,789
Interactions, supervision/discipline	-12.043*** (3.610)	5035
Interactions, teachers	-5.762*** (2.035)	42,447
<i>Panel B: Additional Child Care Program Features</i>		
Parents interactions with management	-7.958*** (2.681)	26,445
Learning and academics	-5.351** (2.140)	35,718
Referrals and recommendations	-5.822** (2.665)	21,944
Search and facility visit	-5.440* (2.899)	20,601
Program cost and fees	-13.145*** (4.242)	6,100
Accessibility	-8.747* (4.779)	6,259
Regulation features	-10.876*** (3.061)	10,758
Religious affiliation	-9.108 (6.592)	895

Notes: Each cell presents the coefficient on the log of county median household income and its standard error (in parentheses) clustered at the county-level. Each coefficient and standard error come from a separate regression. All models include the full set of controls listed in Appendix Table 1, calendar quarter dummies, and county fixed effects. *, **, and *** indicate that a given coefficient is statistically significant at the 0.10, 0.05, and 0.01 level, respectively.

[Restaurants](#)
[Nightlife](#)
[Home Services](#)
[Write a Review](#)
[Events](#)
[Talk](#)
[Log In](#)

WhizKidz Preschool

Claimed

★ ★ ★ ★ ★
20 reviews
[Details](#)

[Write a Review](#)
[Add Photo](#)
[Share](#)
[Bookmark](#)

Preschools, Child Care & Day Care

7214 East Jenan Dr
Scottsdale, AZ 85260

[Get Directions](#)

(480) 999-0944

[whizkidzpreschool.com](#)

[Send to your Phone](#)

[Watch video](#)

Photo of WhizKidz Preschool, Scottsdale, AZ, United States

[See all 25](#)

[Schedule A Tour!](#)
[Contact Us](#)

Today 6:30 am - 6:00 pm Open now

Hours

Mon	6:30 am - 6:00 pm
Tue	6:30 am - 6:00 pm
Wed	6:30 am - 6:00 pm
Thu	6:30 am - 6:00 pm
Fri	6:30 am - 6:00 pm Open now
Sat	Closed
Sun	Closed

[Edit business info](#)

More business info

Accepts Credit Cards **Yes**

Accepts Apple Pay **No**

Accepts Bitcoin **No**

"The campus is beautiful with a **huge outdoor play space** with a soccer field and garden." in 1 reviews

"They have quickly become a **very important part** of our family." in 1 reviews

"We originally started him at another center but sought an alternative following some issues with the feeding schedule and **constant staff changes**." in 1 reviews

[Ask the Community](#)

Appendix Figure 1: Sample Yelp Child Care Business Page



Jennifer T.

Chicago, IL

👤 104 friends

★ 20 reviews

[Share review](#)

[Embed review](#)

[Compliment](#)

[Send message](#)

[Follow Jennifer T.](#)

★★★★★ 7/2/2015

I want to echo all of the other reviews here. I was really nervous about day care, but this place is top notch. As mentioned, Suzy is fantastic and makes me feel my daughter (15 months when she started) is truly cared for. I ran into some parents from my daughters class, and I feel they summed it up - They love my daughter, and we love them! The staff are fantastic and truly diverse. The children are very diverse as well. This is very important to me. My daughter is used to having lots of family give her much attention, and the transition took only a week before our daughter was comfortable and looking forward to seeing everyone.

I love that they have stroller parking, park visits, flower bombs in vacant lots, weekly newsletters with pictures and description of what they were doing so we can reinforce. She comes home with presents for mom and dad, craft projects and I wish I had the energy and patience to do this at home. Our daughter was so much more advanced that we gave her credit for and we learned this by bringing her to Willow. She minds them much better, lol! She sits at a table and eats food as a group with other children, washes her hands, and doesn't tear the house down when diaper changed. So it was nice to see what she was capable of and where we were holding her back. She is interested in so much more (letters), different books, songs, etc.

What was really important to me was flexibility (half days) and understanding of different food requirements. There are vegetarian options for kids, and the food is organic and great. We only do meat, fruit and veggies, and this has worked well, having to supplement meat sometimes when things can't be separated, but that has been super easy. They really make it work. Suzy will scramble her an egg if need be. They ALWAYS let me know how she naps, eats, and how her day was. Sometimes, she needs to be close to one of them and that is important that they recognize it and help her. The older classrooms look awesome. I also like the variety of teachers. They mix it up, meaning there are different teachers at different times with different groups, which has to help their fatigue and patience.

There is GOOD energy here and they really do love the children. They are stimulated, they play and I look forward to bringing my daughter here. It seems like it would be a cool place to work, too. They just opened infant part, which is awesome!

6 people voted for this review



Useful 7



Funny



Cool 1

Appendix Figure 2: Sample Yelp Consumer Review

Appendix Table 1: Summary Statistics for the Analysis Variables

	Full Sample	Bottom Income Decile	Top Income Decile
<i>Panel A: Consumer/Review Characteristics</i>			
Number of friends (no.)	27.03 (117.59)	38.18 (158.09)	20.77 (92.69)
Number of Yelp reviews (no.)	26.40 (100.39)	40.72 (130.14)	19.48 (58.67)
Profile picture of user (%)	0.561 (0.496)	0.594 (0.491)	0.527 (0.499)
Review includes a photograph (%)	0.039 (0.193)	0.042 (0.200)	0.040 (0.195)
Review rated “useful” by others (%)	0.646 (0.478)	0.754 (0.430)	0.528 (0.499)
Dictionary words (no.)	90.91 (4.22)	91.11 (4.27)	90.41 (4.16)
Informal speech (%)	0.495 (0.500)	0.476 (0.499)	0.503 (0.500)
Swear words (%)	0.021 (0.144)	0.029 (0.168)	0.011 (0.105)
Netspeak (%)	0.152 (0.359)	0.143 (0.350)	0.157 (0.364)
Filler words (%)	0.009 (0.096)	0.009 (0.093)	0.009 (0.094)
Assent (%)	0.213 (0.409)	0.212 (0.409)	0.200 (0.400)
<i>Panel B: Firm Characteristics</i>			
Claimed business (%)	0.826 (0.379)	0.740 (0.439)	0.870 (0.337)
First Yelp review (%)	0.053 (0.225)	0.115 (0.320)	0.021 (0.144)
<i>Panel C: County Characteristics</i>			
Unemployment rate (%)	5.754 (2.132)	8.340 (2.560)	3.518 (0.503)
Median household income (%)	70,890 (17,637)	49,478 (4,103)	106,966 (4,452)
Observations	48,630	5,144	4,783