Who Benefits From Child Care Ratings? Evidence From Minnesota's ParentAware Program

Jonathan Borowsky*

This version: January, 2019

Abstract

This paper uses geocoded data on child care enrollments to estimate the parameters of a discrete choice model of child care demand, and from those parameters estimates the direct value of the ratings to consumers of child care. I specify a model of the Berry et al. [1995] type, where consumers vary in their tastes, income, and geographic location. I expect that ratings are endogenous. I follow an intuitive strategy to measure the treatment effect of the ratings using panel data that includes observations of the rated providers before they are rated. I apply results from Train [2015] to present closed-form expressions for the value of the ratings to each type of consumer. I estimate the yearly direct value of the ratings to be more than $3 million, more than justifying the cost of creating and administering ParentAware.

*Department of Economics, University of Minnesota; borow080@umn.edu. Thanks to Thomas Holmes, Elizabeth Davis, Amil Petrin, Aaron Sojourner, Aradyha Sood, and Joel Waldfogel for support, guidance, and feedback, and additional thanks to Davis and Sojourner for providing access to the data. Thanks to participants in the University of Minnesota Applied Micro Workshop for their comments. For now, the results should be considered preliminary. The errors and flaws are my own.

1 Introduction

Encouraged by federal programs, at least 42 states have implemented child care Quality Rating and Improvement Systems, (QRIS) which give providers the opportunity to be evaluated by regulators and assigned a quality rating. This paper uses geocoded panel data on child care providers and consumers to measure the impacts of one such program, Minnesota's ParentAware ratings. The ParentAware program is in large part an informational intervention, designed to improve the quality of child care by creating quality information and distributing it to consumers. I find that the ratings provide a yearly direct value to consumers of more than $3 million dollars, more than justifying the cost of creating and administering ParentAware.

If parents value quality child care but have difficulty discerning high from low quality providers, and if high quality is more expensive to produce, the child care market may suffer from a lemons problem which leads the market to undersupply high quality care. Given the body of evidence linking early childhood experiences to social and intellectual development, improving the quality of child care is a priority for policymakers concerned with inequality and opportunity. Crude or poorly designed measures can perversely distort provider incentives, however systems that rely on detailed assessment and measurement are expensive. For example, ParentAware involves qualified assessors conducting a site visit and administering a scale that measures the quality of the developmental environment, as well as many other measures. According to the Minnesota Office of the Legislative Auditor, this program cost more than $9 million to implement. It is therefore imperative to be able to measure the effects of these programs and assess whether the value that they provide to households is worth the cost. My estimates suggest that, in the case of ParentAware, the benefits are worth the costs.

It is also important to understand heterogeneity in the benefits that informational interventions provide. *Ex ante*, the effect of quality ratings on socioeconomic disparities in child care quality is ambiguous and depends of the relative magnitudes of different effects. The benefit from programs like ParentAware depends on the marginal impact of the provided information on choices, and the amount of benefit will be larger for consumers who have a higher chance of their choice being affected by the additional information. The benefit from the ratings will thus be higher for consumers whose choice set includes many providers whose quality levels are varied and lower for consumers with few choices or whose choices have less variation in quality. On the one hand, less privileged households may be more likely to live in denser urban areas where they can access a range of providers, and may have more difficulty independently discerning quality levels. On the other hand, higher quality providers may be concentrated in affluent neighborhoods and less privileged households cannot be induced to use high quality providers if none are accessible to them. The value of the ratings to households varies geographically and depends on the role of travel costs in household decision-making and the spatial distribution of providers of different types. In order to know how the value of the ratings to different consumers in different areas, we need a model of who shops where.

This paper will estimate the value of Minnesota's ParentAware ratings to households in different locations. I use panel data on provider enrollments, prices, and characteristics, collected by Child Care Aware of Minnesota, a non-profit that surveys all Minnesota licensed child care providers in order to provide data to child care resource and referral agencies, and assembled by Davis et al. [2018] to study geographic variation in child care access. Using this data I estimate a mixed logit model of child care provider choice of the Berry et al. [1995] type. While the classic application of this method is to implement a model with random coefficients on the taste for product characteristics, it also provides a powerful tool for modeling service markets where location is important to choice. A random coefficient can be thought of as a multiplicative interaction between the product characteristic and a latent taste variable. Travel costs are also an interaction, whose form is not multiplicative but is a function of the distance between the provider and the household. The parameters of a travel cost function can be estimated in the same way as the variance parameters of a random coefficients model. The result is a demand model that accounts for spatial substitution patterns where nearby providers compete more intensely with one another than distant ones.

This paper is not the first to use a structural model of demand to calculate the value of an informational intervention. My approach is similar to that used by Jin and Sorensen [2006] to determine the value of National Center for Quality Assurance health plan ratings. Nor is it the first paper to assess informational interventions in the child care setting. Xiao [2010] uses a structural model of child care demand to evaluate the National Association for the Education of Young Children child care accreditation system, finding modest welfare benefits from accreditation system.[1] Herbst [2016] examines the impacts of QRIS systems using national and a difference-in-differences strategy based on differential timing in QRIS adoption. Among other results, he finds that QRIS adoption causes an increase in child care employment, which can be taken as a proxy for child care utilization, and an increase in the educational credentials of people employed in child care, which can be taken as a proxy for quality levels.

In the context of the existing literature, this paper makes two contributions.

First, using geocoded provider data and the Berry et al. [1995] estimation framework, I solve the thorny market definition question that must be addressed in any rigorous study of the child care market. People tend to use child care that is close to their home. Estimates from the National Survey of Early Care and Education, conducted in 2016 by the federal Department of Health and Human Services, suggests the average distance between home and provider, among children using a center-based child care provider, is 4.6 miles for children 0-3 and 3.9 miles for children 4-5, a small distance compared to the size of most population centers. (National Survey of Early Care and Education Project Team [2016]). As a result, child care "markets" are overlapping rather than discrete. Most likely due to data limitations, previous studies like Xiao [2010] have defined child care markets using statisti-

[1] The relationship between this paper and Xiao [2010] is discussed in greater detail in the literature review section below.

cal areas such as ZIP codes. Since child care providers may be distributed unevenly across close-together statistical regions, using statistical regions as a proxy for child care markets risks providing an inaccurate picture of child care access and variety in different locations. This is particularly problematic since the diameter of ZIP codes and other statistical areas varies in systematic ways, making it likely that any measurement error is correlated with neighborhood characteristics that may also be related to child care demand. In this paper, I estimate demand in a single statewide market with travel costs, replacing assumptions about market boundaries with assumptions about the structure of travel costs. The parameters of the travel cost function are estimated within the model. I propose that this approach provides a much more flexible and realistic picture of the child care market than would be possible with arbitrary market boundaries.

Second, I explore variation in the welfare benefits provided by the quality ratings system. As I have argued above, the benefits of quality ratings are inherently heterogenous, as the probability that ratings information will affect the choice of provider depends on the characteristics of the available choice set. I model household heterogeneity using demographic information from the 2011-15 American Consumer Survey (ACS). Each of the more than four thousand Minnesota census block groups that, in the ACS estimates, includes at least one child age 0-5 is a different consumer type with a location at the population centroid of the census block group and a weight equal to the number of children 0-5. Not only is this detailed cross-sectional demand variation helpful for estimating the model, it also means that I am able to conduct welfare analysis at the block group level and use the model to estimate regional variation in the benefits from the quality ratings.

My approach for assigning a money value to the ratings is like that of Jin and Sorensen [2006]. That is to say, I compare choices with the ratings to the model's counterfactual choices for a world without the ratings information, both evaluated according to the utility function of a consumer informed by the ratings information. The question in the welfare counterfactual is "what would a consumer informed by the ratings need to be paid in order to have their decision made by a consumer with identical preferences, who did not have access to the ratings." I make an incremental addition to their method by noting that it is a special case of the framework that Train [2015] derives for welfare calculations when a consumer makes choices based on incomplete information. This provides me with closed-form expressions for the welfare quantities, rather than needing to simulate choice draws.

There are two critical problems of endogeneity that must be resolved in order to use a structural model to measure the value of quality ratings, especially in a market as rich with product variation as the child care market.

First, as Jin and Sorensen [2006], Xiao [2010], and Dranove and Jin [2010] all emphasize, we must expect that quality ratings are endogenous. After all, they are an attempt to measure product quality, and it's reasonable to expect them to be correlated with other quality information that consumers may observe, for example through advertising or provider reputation. Any econometric specification that ignores this will systematically overestimate the impact of quality ratings. In their influential study of restaurant hygiene score cards, Jin and Leslie [2003] use an intuitive panel data strategy to address this, comparing the revenue of highly rated restaurants before and after their rating is disclosed. I use a similar strategy, controlling in the demand model for the time-invariant component of provider quality that is observed by households, but unobserved by the researcher and associated with the rating ultimately assigned.

Second, I expect that price will also be endogenous, as it will be in any model of a product market where consumers have access to information about product quality or characteristics that is not present in the data available to the researcher. I address this using an adaptation of the instruments strategy used by Berry et al. [1995]. For each provider, I construct price instruments based on a distance-weighted sum of the characteristics of nearby competing providers, capturing the expected inverse relationship between markup and local competition.

In section 2 of this paper, I survey the empirical literature on quality ratings and other informational interventions. In section 3, I introduce the data that will be used to estimate the model. In section 4, I present the economic model. This section has two components. First is the model of choice of child care provider that will be estimated, which is a standard logit discrete choice model of demand, with allowance for household heterogeneity based on varying demographics and based on unobserved variation in tastes. Second is an exposition of the method for calculating welfare quantities, following Train [2015]. In section 5, I discuss the estimation methodology. I use the method of Berry et al. [1995] to estimate the demand model by the Generalized Method of Moments. The inversion of Berry [1994] is used to linearize the model and enable the use of instruments for price. A modified version of the Berry et al. [1995] instruments are used, where the instruments are weighted averages of the characteristics of nearby competitors. In section 6 I give a detailed presentation of the results from estimation, which show that the 4 star rating is positive, and other ratings are negative. Overall the ratings are valuable to consumers.

2 Literature Review

This paper contributes to the literature on the effects of quality ratings or "score cards", much of which is reviewed in Dranove and Jin [2010]. The existing research suggests that disclosure of product attributes can have a big impact on consumer choice. Jin and Leslie [2003] consider Los Angeles County's introduction of a rule requiring restaurants to post a letter grade that reflects their performance on a health inspection. They find substantial evidence that this requirement led to an improvement in the hygiene performance of affected restaurants. Bollinger et al. [2011] study the impact of New York City rules that require restaurants to post calorie values of menu items on sales in Starbucks stores, finding that these these rules led consumers to choose lower-calorie foods and increased the sales of food in Starbucks establishments that were close to competing Dunkin Donuts stores. Disclosure rules are thus an attractive intervention for policy-makers.

Evidence on the effectiveness of *voluntary* disclosure regimes is more mixed. In their study, Jin and Leslie [2003] compare the effectiveness of the mandatory letter grade regime to a transitional period when some municipalities in Los Angeles County required the letter grades to be posted, but others did not, characterizing the latter regime as voluntary disclosure. They find that the effects of the voluntary disclosure regime are much less. Similarly, Mathios [2000] studies the effect of the Nutrition Labeling and Education Act, which mandates disclosure of nutritional information, on the sales of different types of salad dressing. This paper compares the salad dressing market before the introduction of NLEA, when some products displayed nutritional information labels to a period after the introduction of NLEA, when such labeling became mandatory, and finds substantial differences in consumer behavior. Hotz and Xiao [2013] provide a theoretical treatment that illustrates conditions where firms choose not to participate in quality disclosure for strategic reasons related to the effect of quality disclosure on markups through changed competition patterns.

A close comparison can be made between the present study and Jin and Sorensen [2006]. That paper examines the National Center for Quality Assurance ratings, which are a voluntary ratings system for health plans. Consumers may be able, at least to some extent, to discern quality without the availability of the ratings, and in that paper, the authors use non-public ratings of health plans to separately identify the treatment effect of the ratings from the ratings' correlation with already-available information about quality. The present approaches the same problem by using data on the period before the ratings are available to control for the already-known information about quality that is likely to be correlated with the ratings.

A close comparison can also be made between the present study and Xiao [2010]. That paper examines the privately administered system of child care center accreditation managed by the National Association for the Education of Young Children (NAEYC). Like the present study, that paper estimates the value to consumers of the ratings in a discrete choice framework. The present study differs from Xiao's work in three ways. First, I am able to use true panel data, consisting of multiple observations over time of the child care providers, including data on the enrollment of rated centers before they were rated. My strategy for controlling for endogeneity in the ratings, using these pre-rating observations, is more direct than Xiao's instrumental variables strategy. Second, I model consumer heterogeneity.

On the other hand, Xiao [2010] endogenizes the weight placed by consumers on the ratings. In her model agents have two different signals of quality, one coming from the ratings, and one that is "reputation" independent of the ratings. The longer a provider has existed the more informative the "reputation" signal is. The weight placed by the agent on each signal depends on the informativeness of that signal, and hence in Xiao's model the effect of ratings is moderated by the operational age of the provider, with ratings making the biggest difference for providers that are more recent entrants. I do not yet have data on how long the providers in my sample have existed, but expect to be able to include the relationship between operational age and quality ratings in future versions of this paper.

Jin [2005] considers the strategic incentives to disclose quality information, concluding that HMOs use participation in National Center for Quality Assurance ratings to distinguish themselves from competitors, and in the ratings in highly competitive markets are less likely to participate in the ratings.

An interesting thread in the literature deals with heterogeneity in consumers' use of disclosed product information. One way to put it is that not all customers are listening. This is of particular interest in the child care quality setting because of policymakers' interest in the choices made by low income families. Milyo and Waldfogel [1999] do not study quality disclosure, but rather price disclosure in the form of price advertising. Using a difference-in-differences strategy made possible by the *44 Liquormart* case, which legalized liquor store advertising in Rhode Island, they find limited effects of advertising on prices. The subset of stores who advertise after legalization cut some prices, but only on the products that they advertise or those advertised by competitors.

3 Data

3.1 Provider Panel

I use a unique panel data set based on an annual census of Minnesota licensed child care providers. Child Care Aware of Minnesota maintains a database of child care providers as part of NACCRAware, a national child care data system designed for use by child care referral agencies. Providers are surveyed annually by Child Care Aware of Minnesota, on a rolling basis, in order to keep this database up to date. Information is self-reported by the providers, and includes enrollment numbers by age group, price by age group, quality rating information, as well as several other provider characteristics such as accreditations and nonprofit status. Davis et al. [2018] have prepared panel data on Minnesota child care providers based on periodic pulls from this database, merged with additional information from state government sources and licensing records. Provider location is geocoded based on provider address from the Child Care Aware data and licensing records. The panel data covers fiscal years 2012-2015 and includes information about child care centers, family day cares, and certain public child care programs, such as Head Start and pre-K programs in schools. This paper focuses on licensed child care centers.

Table 1: Provider Descriptive Statistics, Fiscal Year 2015

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Enrollment	1,194	56.861	36.649	4.000	34.085	68.500	392.667
Weekly Price	1,194	217.946	58.284	12.084	181.925	253.198	441.990
Licensed Capacity	1,194	80.214	59.214	0	43	108	1,140
USDA Food Prog.	1,194	0.330	0.470	0	0	1	1
Non-Profit	1,194	0.366	0.482	0	0	1	1
Accreditation	1,194	0.315	0.465	0	0	1	1
Rating Stars	1,194	1.328	1.828	0	0	4	4

3.2 Supplementary Data

Provider data is supplemented by block group level household demographic data from the American Community Survey 2011-15 estimates, accessed through the NHGIS geocoded census data system. Household "locations" are actually the block group centroids. Each location is weighted by the ACS estimates of the number of children 0-5 in that block group. "Percent low income" is fraction of population in the block group under 200% of the poverty level. "Percent college" is fraction of adult population in the block group with bachelors degree or greater education. Household-provider distances are straight-line distance between the block group centroid and the provider address, calculated using Vincenty's formula.

4 Model

4.1 Demand

I represent choice of child care provider using a standard discrete choice model of product choice. Specifically, consider household i's decision over what child care arrangement to use. Suppose that i can choose any provider within R miles, where R is a distance radius around household i's location that is chosen to be large compared to child care travel distances, such as 50 miles. Then, the choice set for household i in period t is

$$J_{it} = \{j \in J_t | d_{ij} < R\} \cup \{0\}$$

where d_{ij} is distance between household i and provider j, and choice 0 is an outside option. The outside option represents any choice not explicitly represented in the choice set. Here, that would include parental care, care by friends and family, care in a licensed family day care, or care in a publicly provided center such as a Head Start program.

Following Berry, Levinsohn, and Pakes (1995) (hereinafter, "BLP",) I allow for a choice-specific utility that depends on both provider characteristics and characteristics (observed or postulated) of households.

$$u_{ijt} = X_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \gamma d_{ij} + \sum_h \sigma_h x_{jht}\nu_{ih} + \epsilon_{ijt}$$

Here, ξ_{jt} is a "structural" error term representing unobserved information about provider j in period t that is relevant to all households, and ϵ_{ijt} is error term that is "idiosyncratic" to the household. Choice-specific household utility depends on household "type" in two ways. First, random coefficients may be implemented through ν_{ih}, a random draw from a unit normal distribution, so that consumer i's "taste" for characteristic h is distributed $N(\beta_h, \sigma_h)$.[2] Second, through the effect of the household-provider distance term d_{ij}.

The purpose of allowing choice-specific utility to be different for different "types" of household is twofold. First, it allows for more patterns of substitutability between providers that are more complex than what could be represented in a non-mixed specification. Second, it allows for the model to incorporate important information about households that affects those substitution patterns.

Here it is worth saying a little bit about the role of household-provider distance in the model. Because provider market shares are observed only at the aggregate level, we do not directly observe the distance between households and the providers they choose. Households are assigned to providers endogenously through the demand model. However, the inclusion of household-provider distance allows the model to treat providers that are near to one another as closer substitutes than providers that are far away from one another, in a way that is shaped by specific information on where households live. In this way, the model's treatment of household-provider distance is analogous to the way BLP treat household income.

In the discrete choice model, each household chooses the provider in their choice set that provides the highest choice specific utility. Assuming that the idiosyncratic error term ϵ_{ijt} has the extreme value distribution, and is i.i.d., and normalizing the utility value of the outside option to be centered around zero, the probability that household i will choose provider j is

[2] This version of the paper, however, does not report the results of any random coefficients specifications.

given by the logit choice probability formula.

$$P_{ijt} = \frac{e^{X_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \gamma d_{ij} + \sum_h \sigma_h x_{jht} \nu_{ih}}}{1 + \sum_{k \in J_i} e^{X_{kt}\beta - \alpha p_{kt} + \xi_{kt} + \gamma d_{ik} + \sum_h \sigma_h x_{kht} \nu_{ih}}}$$

Demand, stated as market shares for each of the providers, has the form

$$s_{jt} = \int_i \frac{e^{X_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \gamma d_{ij} + \sum_h \sigma_h x_{jht} \nu_{ih}}}{1 + \sum_{k \in J_i} e^{X_{kt}\beta - \alpha p_{kt} + \xi_{kt} + \gamma d_{ik} + \sum_h \sigma_h x_{kht} \nu_{ih}}} w_i di$$

where w_i is a weight function capturing the proportion of households of each type, based on the proportion of households at each location and the assumed distribution of the random taste parameter.

4.2 Calculation of Welfare Quantities

Ratings are not quite like other product characteristics, when it comes to calculating welfare quantities. For illustration, it may be helpful to consider the hypothetical of the low-rated provider's loyal client. Suppose that there is a provider whose quality rating indicates a quality level lower than what consumers would have expected without the rating. Further suppose that there are some households who choose that provider despite the low rating. (Perhaps the cost is also low, or the household receives a good idiosyncratic utility draw for that provider.) We now ask how well that consumer would be if the quality information was not available. For this consumer, the quality information cannot have been marginal. If the chosen provider is the best choice even with the negative quality information, it will also be the best choice without the negative quality information. Intuitively it is clear that this consumer is *unaffected* by the quality ratings.

This issue is recognized in the existing literature. Jin and Sorensen [2006] argue that to properly determine the utility value of ratings, the value of the counterfactual state where the ratings exist should be computed using the *decision rule* determined by removing the ratings from the choice value function, but *utility values* determined using the ratings. One way to think about this method is that the compensating variation for the ratings is the amount of money that a consumer *with* the ratings information would need to receive in order to compensate for allowing their decision to be made by a consumer, with identical preferences, who didn't have access to the ratings. Jin and Sorensen [2006] use a simulation method to compute the value of the without-ratings decision rule under the with-ratings utility values. In this section, I develop closed-form expressions for the same quantity by applying the framework of Train [2015] for welfare calculations when experienced utility differs from anticipated utility.

Now I will describe how to calculate the utility value of the ratings. For simplicity, consider a single type of household, with choice set J. Let ζ_j be the non-idiosyncratic portion of the value of choosing j, so that $u_j = \zeta_j + \varepsilon_j$. However, we will suppose that the household doesn't get to see the actual values of ζ_j when they choose. Instead they have a belief $\zeta_j^\omega = \zeta_j + D_j^\omega$ about the expected value of ζ_j. This belief depends in the available information, so that the beliefs with ratings available, ζ_j^B are different from the beliefs with ratings unavailable, ζ_j^A. Let $D_{BA} = \zeta_j^B - \zeta_j^A = D_j^B - D_j^A$.

Following Train [2015], with some adaptations of notation, the expected utility under information state ω is [3]

$$\log\left[\sum_{j\in J} e^{\zeta_j + D_j^\omega}\right] - \sum_j P_j^\omega D_j^\omega$$

In order to calculate compensating variation, we need to know the utility value of the choices that would be made under information regime A, evaluated under the utility beliefs of information regime B. The difference between the utility beliefs of regime B and regime A is D_{BA}, and so this quantity is

$$\log\left[\sum_{j\in J} e^{\zeta_j^B}\right] - \log\left[\sum_{j\in J} e^{\zeta_j^A}\right] + \sum_j P_j^A D_{BA}$$

[3] The expression can be derived by writing $\zeta_j = \zeta_j^\omega + D^\omega$ and then following the standard derivation of the expected value of the optimal choice in the logit choice model.

This expression can be understood as the anticipated utility under regime B, minus the anticipated utility under regime A, plus an adjustment for the "surprises" when decisions made under regime A turn out to be better or worse than expected when adjusting to the utility beliefs of regime B.

5 Estimation

5.1 Instruments

As is usual in a model of demand for differentiated products, we expect that price is likely to be endogenous. There are many inputs that we do not directly observe that would be expected to determine quality or other dimensions of desirability of child care providers to consumers, and it is reasonable to suppose that the "high-quality" providers should also be higher-priced.

Let x_j^x be a provider characteristics. A price instrument z_j^c is constructed by taking the distance-weighted sum of x_k^c across other providers in the market.

$$z_j^c = \sum_{j \in J_t, k \neq j} \frac{x_k^c}{d_{jk}}$$

Five instruments are constructed in this manner. Instruments are constructed from provider characteristics licensed capacity, nonprofit status, accreditation status, and ParentAware rating status. An instrument is also constructed using a constant for x_k^c, which provides a measure of the density of other nearby centers.

One test of instruments is to examine the results of the "first stage" regression of the endogenous variables on the instruments. If the instruments do not have explanatory power, they cannot be very satisfactory. The specification uses year fixed effects to control for time trends in child care prices. The instruments created from competitor count, accreditation, and ParentAware rated status all have the expected sign and are statistically significant.

The instrument created from non-profit status has the correct sign but is not statistically significant. Somewhat oddly, the instrument created from total licensed capacity is statistically significant but with the opposite sign from that expected. That is, providers that are nearby to other providers with high capacity tend, themselves, to have higher prices. Taken as a whole, the instruments provide substantial additional explanatory power. An F-test of the restriction dropping the five instruments from the model is rejected at the 1% significance level, with an F statistic of 558.76.

5.2 Berry Inversion and Estimation Strategy

I follow the strategy devised by Berry (1994) for how to use instruments to estimate a discrete choice model. The strategy involves inverting the market shares function in order to get a linear problem.

In order to explain how this works, it is helpful to re-write the choice-specific utility function as the sum of three terms

$$u_{ijt} = \delta_j + \mu_{ij} + \epsilon_{ij}$$

The first term, δ_j, captures the purely "vertical" dimension of differentiation between providers. That is, δ_j is everything about provider j that is valued the same by all households, including the unobserved quality term ξ_j, and the value placed on j's observed characteristics by a consumer with average tastes. The second term, μ_{ij} captures the "horizontal" dimension of differentiation between providers; those parts of the household's valuation that depend on the household's type. This might include the effect on choice of the distance between household and provider. Even if all households place the same value on distance, the values of d_{ij} will be different for households at different locations. μ_{ij} will also include the effects of taste variation as expressed through random coefficients. Finally, ϵ_{ij} is the random idiosyncratic error term.

Given that μ_{ij} is a function of the data and some unknown parameters θ, and given a vector of values, $\delta = \{\delta_j\}$, fitted market shares can be computed, conditional on θ and $\{\delta_j\}$, using the formula

$$s_j(\delta;\theta) = \sum_i w_i \frac{e^{\delta_j + \mu_{ij}}}{1 + \sum_k e^{\delta_k + \mu_{ik}}}$$

The strategy for estimating the model has three parts. First, given a candidate value for θ, determine the vector of values $\delta(\theta)$ that matches the fitted market shares $s_j(\delta;\theta)$ to the observed market shares in the data. Second, considering the expression,

$$\delta_j = X_j\beta - \alpha p_j + \xi_j$$

estimate the vector of structural errors $\xi(\theta)$ as the residual of an instrumental variables regression with $\delta(\theta)$ as the left hand side. Third, using this vector of structural errors as an input into a GMM objective function, and re-computing δ and ξ for each candidate value of θ find the value of θ that minimizes that objective function. In the current version of the analysis, the only moment condition is the one from the demand equations, $E(\xi_j|Z_j) = 0$. The objective function I use to estimate θ is thus relatively simple.

$$\hat{\theta} = \arg\min_{\theta} \xi(\theta)^T \xi(\theta)$$

5.3 The BLP contraction

The foregoing discussion assumes that there is a method of determining the vector δ that matches the fitted market shares $s(\delta;\theta)$ to the observed market shares s^0. In order to do so I use the contraction described by BLP. This method uses a process of iterative adjustment. Define the operator $T\cdot$ by

$$T \cdot \delta_j = \delta_j + \ln s_j^0 - \ln s_j(\delta;\theta)$$

BLP show that this operator is a contraction and thus that it has a unique fixed point, which can be found by iteratively applying it to an initial "guess". Since the fixed point occurs when $s_j(\delta;\theta) = \ln s_j^0$, this provides a computational method for calculating $\delta(\theta)$. An initial value of δ is set, and then the contraction iteratively applied until the differences between δ and $T \cdot \delta$ are small compared to a specified tolerance.

5.4 Identifying a Treatment Effect

In specifying the model, the ParentAware star ratings are included in the provider characteristics matrix X_{jt}. We should, however, be concerned that the demand unobservable ξ_{jt} will be correlated with the ParentAware ratings variables. ξ_{jt} captures whatever information about providers that is unobserved by the researcher but that households know and incorporate into their choice decisions. If, as seems reasonable to expect, the providers that have high values of ξ_{jt} in the data generating process are more likely to be assigned a four star rating, then the coefficient on the four star rating parameter may be biased upwards.

In order to address this, I follow a difference-in-differences strategy. Implicit in this strategy is the assumption that the component of ξ_{jt} that is correlated with the eventual rating is stable over time. Each provider is assigned to a "group" based on the highest ParentAware rating they receive. Thus, a provider whose highest star rating is four stars is assigned to the four stars group. The providers that are never rated in the data are their own group. A dummy variable is assigned to each group and included in X_{jt}. These dummy variables account for the average differences between the providers that receive, for example, four stars, when they are rated, and those that receive three stars, or that never choose to become rated.

As part of this difference-in-differences strategy, I also include year fixed effects by incorporating year dummies into the characteristics matrix X_{jt}. Year fixed effects are equivalent to allowing the value of the outside option to be different in different years. This is necessary because the overall demand for child care, and for center-based care, is not constant over time, it is increasing. Since the likelihood of being rated is correlated with time – the ratings are only present in the later years – if we did not account for this overall demand trend it might bias the ratings coefficients upwards. An alternative to using year fixed effects would be including a linear or quadratic time trend.

6 Results

6.1 Baseline Demand Models

In this subsection, I compare the estimates from the mixed logit demand model to logit models that do not account for the role of geography and travel costs in child care demand.

Table 2 shows these results. In all three specifications, the outcome variable can be understood as a score that captures the utility value to an average consumer of choosing product j. For models A and B the outcome variable is defined as $\ln s_j - \ln s_0$, the linearization of the logit model that Berry [1994] suggests in order to allow instrumental variables to be employed. Coefficients for models A and B are by ordinary (OLS) and instrumental variables (2SLS) least squares respectively. Model C is a mixed logit model where the household type varies with location, with a single household-product utility interaction term, a linear travel cost. Model C is estimated using nonlinear two stage least squares using the Berry et al. [1995] estimation algorithm.[4] The "outcome" variable in a mixed logit model is a mean utility score that is most intuitively understood by noting that if the utility specification contains no household-product interaction terms then the model specializes to the $\ln s_j - \ln s_0$ model of A and B. Thus if the model C estimation algorithm is run with the coefficient on travel distance dropped (or equivalently, fixed at zero), the results are the same as for model B. As a consequence of this, the coefficients in all three columns are on the same scale and can be compared directly.

[4] That is to say, I use BLP's nonlinear GMM estimation algorithm, but choose the weighting matrix that specializes the GMM estimator into 2SLS.

All of these specifications include year fixed effects to control for population growth and statewide trends in the demand for child care.

Table 2: Baseline Models

	OLS	2SLS	Nonlinear 2SLS /BLP
	(A)	(B)	(C)
Distance (Miles)	.	.	-0.093^{***}
			(0.002)
Price (Weekly)	0.001^{***}	-0.00004	-0.035^{***}
	(0.0002)	(0.0003)	(0.001)
Star Rating – 1	-0.182^{*}	-0.186^{*}	-0.425^{*}
	(0.096)	(0.096)	(0.249)
Star Rating – 2	-0.00001	-0.005	-0.440^{**}
	(0.075)	(0.075)	(0.194)
Star Rating – 3	0.107	0.094	-0.377
	(0.151)	(0.151)	(0.391)
Star Rating – 4	-0.007	-0.002	0.423^{***}
	(0.053)	(0.053)	(0.138)
Group – 1	0.079	0.067	-0.284
	(0.082)	(0.082)	(0.212)
Group – 2	0.059	0.049	-0.117
	(0.065)	(0.065)	(0.168)
Group – 3	0.058	0.052	-0.327
	(0.119)	(0.119)	(0.307)
Group – 4	0.037	0.033	0.131
	(0.052)	(0.052)	(0.134)
Capacity	0.006^{***}	0.006^{***}	0.011^{***}
	(0.0002)	(0.0002)	(0.0005)
USDA Food Program	-0.099^{***}	-0.103^{***}	0.035
	(0.021)	(0.021)	(0.055)
Nonprofit	0.014	0.002	-0.364^{***}
	(0.021)	(0.021)	(0.055)
Accreditation	0.070^{**}	0.099^{***}	0.671^{***}
	(0.030)	(0.033)	(0.085)
Constant	-9.345^{***}	-9.231^{***}	0.497^{***}
	(0.045)	(0.071)	(0.184)
Year F.E.	Yes	Yes	Yes
Observations	3,426	3,426	3,426

Note: *p<0.1; **p<0.05; ***p<0.01

Comparing the price coefficients in Table 2, we can see that price has a coefficient that is positive and statistically significant in model A, which was estimated with ordinary least squares. Taken at face value, the coefficient implies that consumers prefer to spend more rather than less on child care providers. This was to be expected given our presumption that price is correlated with unobserved quality. Model B uses instrumental variables. Here, the price coefficient is small and its difference from zero is not statistically significant. In that model there is no accounting for travel costs, this model treats all providers as equally substitutable with one another regardless of distance. Model C includes travel costs, so the estimates are based on a demand structure where the degree of substitutability depends on geography. In the coefficient estimates for this model, distance and price are both statistically significant and have the expected sign. The relative size of the coefficients, $\theta_{Distance}/\beta_{Price} = 2.64$, multiplying by 52 to obtain a yearly quantity, suggests that a typical consumer is willing to pay about \$137 extra per year to avoid an additional mile of distance between home and the child care provider. This substantial distance penalty is consistent with our expectation that consumers prefer child care that is within a few miles of the home.

The key coefficients of interest are the ones associated with the star ratings. The omitted category is "unrated", and so the ratings coefficients can be understood as differences from the unrated category. In Model C, the coefficient on the highest rating, 4 Stars, is 0.423. Scaling this to the coefficient on price, $\beta_{4Star}/\beta_{Price} = \$12.09$, and multiplying this by 52 implies that a typical consumer is willing to pay about \$628 extra per year in order to use a 4 Star rated provider compared to an unrated provider. On the other hand, the coefficients on the lower rating, 1-3 stars are all negative, and all have values around -0.4. The results imply that providers that receive ratings lower than 4 Stars are perceived as worse than unrated providers. Scaling these coefficients in the same way as above suggests that, for example, a typical consumer would require about \$630 of compensation per year in order to accept a 1 Star rated provider rather than an unrated. All of the coefficients on the star ratings are statistically significant except the one on 3 Stars. The non-significant coefficient on the 3

Star rating may be attributed to the fact that there are comparatively few providers with this rating.

6.2 Models With Demographics

Table 3 shows the results of specifications where household demographics enter into household demand but do not interact with provider characteristics. One way to think about the interpretation of this is that the household demographics affect the likelihood of choosing the outside option, but do not affect the relative attractiveness of the various inside goods.

The parameter estimates resulting from these specifications are a little perplexing. Column D shows a specification that includes the low income variable, defined as the percentage of households in the block group whose income is less than 200% of the poverty level. The coefficient on this demographic variable is large, negative, and statistically significant, which would suggest that demand for center-based child care is in block groups with many low-income families. The estimate on the distance parameter jumps up to -0.497, compared to the value of -0.093 that was estimated for Column C of Table 2. However, while the model with the income demographic has these distinctive results, the models that include the college variable, defined as the percentage of people over 25 in the block group whose education level is at least a bachelors degree, are essentially the same as Column C of Table 2. Column E is the estimates for the model that only includes the college variable. Column F is the model that includes both demographic variables, nesting both Column D and Column E. In both of these cases the estimated coefficients on the demographics are close to zero, and hence the other parameter estimates are almost identical to Column C of Table 2.

I do not have a compelling explanation for why the specification in column D should be so different from the nesting specification in column F. This issue merits further investigation and consideration to ensure that the estimates presented are reliable.

Table 3:

	(D)	(E)	(F)
Distance (Miles)	-0.497***	-0.093***	-0.093***
	(0.032)	(0.008)	(0.023)
% Low Income	-5.242***	.	0.000
	(0.509)		(1.59)
% College	.	0.001	0.000
		(0.310)	(0.569)
Price (Weekly)	−0.034***	−0.035***	−0.035***
	(0.001)	(0.001)	(0.001)
Star Rating – 1	−0.477*	−0.425*	−0.425*
	(0.253)	(0.249)	(0.249)
Star Rating – 2	−0.509***	−0.440**	−0.440**
	(0.197)	(0.194)	(0.194)
Star Rating – 3	−0.425	−0.377	−0.377
	(0.397)	(0.391)	(0.391)
Star Rating – 1	0.452***	0.423***	0.423***
	(0.140)	(0.138)	(0.138)
Group – 1	−0.407*	−0.284	−0.284
	(0.216)	(0.212)	(0.212)
Group – 2	−0.076	−0.117	−0.117
	(0.171)	(0.168)	(0.168)
Group – 3	−0.490	−0.327	−0.327
	(0.312)	(0.307)	(0.307)
Group – 4	0.203	0.131	0.131
	(0.136)	(0.134)	(0.134)
Capacity	0.010***	0.011***	0.011***
	(0.0005)	(0.0005)	(0.0005)
USDA Food Program	−0.031	0.035	0.035
	(0.056)	(0.055)	(0.055)
Nonprofit	−0.326***	−0.364***	−0.364***
	(0.056)	(0.055)	(0.055)
Accreditation	0.525***	0.671***	0.671***
	(0.086)	(0.085)	(0.085)
Constant	4.036***	0.499***	0.498***
	(0.187)	(0.184)	(0.184)
Year F.E.	Yes	Yes	Yes
Observations	3,426	3,426	3,426

Note: *p<0.1; **p<0.05; ***p<0.01

6.3 Welfare Calculations

In this subsection, I report calculations of welfare quantities based on the estimates from the econometric model. All of the welfare calculations are based on the simplest specification of the BLP model, column C from Table 2.

6.3.1 By Year

Table 4: Welfare Quantities

FY	Actual	Counterfactual	Adjustment	Benefit
2014	\$100,975,800.42	\$93,104,316.91	-\$5,768,092.55	\$2,103,390.96
2015	\$115,533,611.15	\$106,222,502.29	-\$6,270,825.54	\$3,040,283.31
2016	\$115,804,854.12	\$107,021,278.21	-\$5,313,216.28	\$3,470,359.63

Table 4 shows the calculation of the total welfare benefit from the ParentAware ratings. The first column, "Actual" shows a money scaling of the total expected utility calculated from the model as follows. First, for each block group type i, I calculate $\frac{52}{\beta_{Price}} \log \sum_j e^{V_{ij}}$, the expected utility from the model to a consumer of that type evaluated over possible values of the idiosyncratic utility draw ε_i, scaled to a money value by dividing by the price coefficient, and translated into a yearly value by multiplying by 52. These per-person values are summed, weighted by the number of children 0-5 in that block group, to give the values shown in the table. To calculate the second column, "Counterfactual", I calculate counterfactual choice utilities by starting with the estimated model and setting the coefficients on the ratings to zero. I then calculate the same log sum calculation, scaling, and summing as for the previous quantity, yielding an estimate of the expected utility from the model to a consumer choosing without the ratings information. However, I wish to calculate the value of the without-ratings choices according to the with-ratings utility values. The third column, "Adjustment", is calculated at the consumer type level as $\sum_j P_{ij} D_{ij}$, where D_{ij} is the difference between choice utility with the ratings coefficients set to zero, and the choice utility with the estimated coefficients; and P_{ij} is the probability that i will choose j in the counterfactual with the ratings coefficients set to zero; scaled and summed in the same way. "Counterfactual" - "Adjustment" gives the total expected value to consumers of their choices without the ratings, evaluated using the with-ratings utilities. "Actual" - "Counterfactual" + "Benefit" gives the compensating variation of the ratings. That is, the amount of money that would compensate consumers who had the ratings for having their choices made by consumers with identical preferences except for the ratings.

The results show a rapid growth in the total value of the ParentAware ratings to consumers, up to a yearly total value of almost \$3.5 million in fiscal year 2016. This reflects the increase in the number of rated providers, and consequently in the number of consumers whose choice is influenced by the ratings.

6.3.2 By Location

The value of the ratings is inherently heterogenous. Consumers whose set of available choices includes a variety of different classifications of providers are more likely to have their choice affected by the ratings. Consumers with few choices, or whose choices had uniform ratings, would be unlikely to have the ratings affect their choice and would gain little from the ratings information.

Table 5 shows the per-child expected benefit per consumer, aggregated to the county level. The table includes sixteen Minnesota counties. The counties shown in the table are all the counties whose share of the children age 0-5 is at least 1%, according to the ACS. The column labeled "Benefit" is calculated by evaluating the yearly expected benefit at the block group level, and then averaging over the block groups in each county, weighted by the number of children age 0-5 in each block group. The column labeled "Benefit Share" is the total expected value of ParentAware to consumers in the county divided by the statewide total, and the column labeled "Child Share" is the number of children age 0-5 in the county divided by the statewide total.

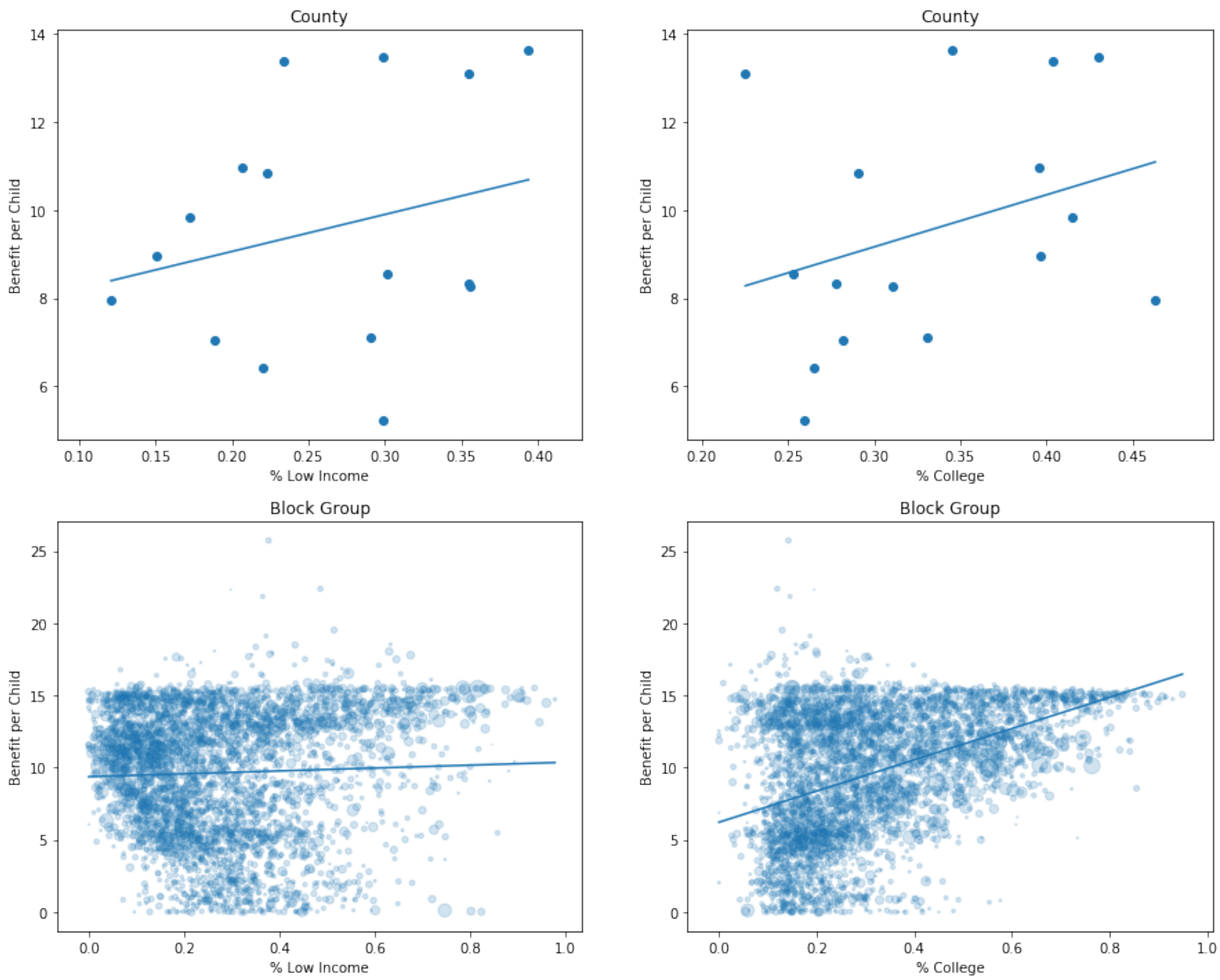
Table 5: Per-person Benefit by County (FY 2016)

County	MSA	Benefit ($)	Child Share (%)	Benefit Share (%)
Ramsey County	Twin Cities	13.64	10.67	14.64
Hennepin County	Twin Cities	13.48	22.65	30.74
Olmsted County	Rochester	13.39	3.06	4.12
Crow Wing County	(No MSA)	13.10	1.07	1.41
Dakota County	Twin Cities	10.96	7.76	8.56
Anoka County	Twin Cities	10.84	6.17	6.73
Washington County	Twin Cities	9.84	4.39	4.35
Scott County	Twin Cities	8.95	2.95	2.65
Stearns County	St. Cloud	8.55	2.75	2.36
St. Louis County	Duluth	8.34	2.99	2.51
Blue Earth County	Mankato	8.27	1.06	0.88
Carver County	Twin Cities	7.97	1.83	1.47
Clay County	Fargo-Moorhead	7.11	1.20	0.86
Wright County	Twin Cities	7.05	2.84	2.01
Sherburne County	Twin Cities	6.41	1.77	1.14
Rice County	(No MSA)	5.24	1.06	0.56

Generally, the benefits of the ratings are greatest in the urban counties. The chart is topped by the counties containing Minnesota's three most populous cities: Minneapolis (Hennepin County), St. Paul (Ramsey County), and Rochester (Olmsted County). This is in line with our expectations that the benefits will be greater in dense cities where consumers have a wider variety of child care choices. The pattern is not absolute, however. Crow Wing, (a central Minnesota county whose principle city is Brainerd) has a larger than usual number of four star centers and consequently a high per-child expected benefit from ParentAware, exceeding that of suburban counties in the Minneapolis-St. Paul metropolitan area.

Figure 1 describes the relationship between the estimated benefits from ParentAware and some demographic characteristics. It should be noted that these estimates are based on the model that includes the distance cost but not any demographics in the choice utility specification, and therefore the variation in local benefits depicted in figure 1 arises wholly from differences in local child care choice sets and provider characteristics, rather than the direct effects of demographics. The top two subplots show county level aggregates, and

Figure 1: Local Benefits of ParentAware, by Demographics



the bottom two subplots show block group level estimates. In the left two subplots, the horizontal axis variable is the percent low income, defined as the proportion of households whose income is less than or equal to 200% of the federal poverty level. In the right two subplots, the horizontal axis variable is the percent college, defined as the proportion of individuals aged 25 and older whose highest education level is at least a bachelors degree. For each of these variables the county level quantity has been computed by taking a weighted average of the block group level variables, with the weight depending on the number of children in the block group (rather than the number of households or adults).

The left group of subplots of figure 1, describing the relationship between income and the benefits of ParentAware, are particularly interesting. At the county level, there is a loose positive relationship between the income variable and the estimated benefits from ParentAware. I interpret this as reflecting the effect of density. Many block-groups with a higher percentage of low-income people are urban block groups that are denser and where there is a greater variety of child care options. Consequently, there is a higher likelihood that the ratings will affect the choice of child care. At the block group level, however, the relationship between the income variable and the benefit variable is weaker. I interpret this as reflecting a balance of the effects of density against the fact that low-income neighbourhoods contain fewer highly rated centers.

The right group of subplots of figure 1 describe the relationship between education and the benefits of ParentAware. They show a positive relationship at both the county and block group level. This is in line with expectations. More college-educated people live in urban areas where the variety of child care options is greater and ratings are more likely to have an effect on choice. Furthermore, areas containing more college-educated people are also more likely to contain highly rated centers.

References

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

Bryan Bollinger, Phillip Leslie, and Alan Sorensen. Calorie posting in chain restaurants. *American Economic Journal: Economic Policy*, 3(1):91–128, 2011.

Elizabeth E. Davis, Won F. Lee, and Aaron Sojourner. Family-centered measures of access to early care and education. *Early Childhood Research Quarterly*, 2018.

David Dranove and Ginger Zhe Jin. Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4):935–63, 2010.

Chris M. Herbst. The impact of quality rating and improvement systems on families' child care choices and the supply of child care labor. IZA Discussion Papers 10383, Institute for the Study of Labor (IZA), 2016.

V Joseph Hotz and Mo Xiao. Strategic information disclosure: The case of multiattribute products with heterogeneous consumers. *Economic Inquiry*, 51(1):865–881, 2013.

Ginger Zhe Jin. Competition and disclosure incentives: An empirical study of hmos. *RAND Journal of Economics*, 36(1):93–112, 2005.

Ginger Zhe Jin and Phillip Leslie. The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451, 2003.

Ginger Zhe Jin and Alan T. Sorensen. Information and consumer choice: The value of publicized health plan ratings. *Journal of Health Economics*, 25:248–275, 2006.

Alan D. Mathios. The impact of mandatory disclosure laws on product choices: An analysis of the salad dressing market". *The Journal of Law and Economics*, 43(2):651–678, 2000.

Jeffrey Milyo and Joel Waldfogel. The effect of price advertising on prices: Evidence in the wake of 44 liquormart. *American Economic Review*, 89(5):1081–1096, 1999.

National Survey of Early Care and Education Project Team. Fact sheet: How far are early care and education arrangements from children's homes? Technical Report OPRE Report No. 2016-10, Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2016.

Kenneth Train. Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples. *Journal of choice modelling*, 16:15–22, 2015.

Mo Xiao. Is quality accreditation effective? evidence from the child care market. *International Journal of Industrial Organization*, 28(6):708–721, 2010.