aawe
wine economics

AMERICAN ASSOCIATION OF WINE ECONOMISTS

AAWE WORKING PAPER
No. 108
Economics

OEXPERT OPINION AND QUALITY PERCEPTION OF CONSUMERS: EVIDENCE FROM NEW YORK CITY RESTAURANTS

Olivier Gergaud, Karl Storchmann
and Vincenzo Verardi

May 2012
ISSN 2166-9112

www.wine-economics.org

Expert Opinion and Quality Perception of Consumers: Evidence from New York City Restaurants

Olivier GERGAUD
(Bordeaux Management School & University of Reims)

Karl STORCHMANN
(New York University)

Vincenzo VERARDI
(Université Libre de Bruxelles & Université de Namur)

Abstract

Exploiting a natural experiment for New York City restaurants we analyze whether consumers' quality perception is influenced by newly appearing expert opinion. As the leading restaurant guide *Zagat* has rated New York City's restaurants since 1979 by drawing on consumer surveys. In 2005, with the first release of the red *Michelin Guide New York City*, *Zagat* faced a serious competition. In contrast to *Zagat*, *Michelin* relies on experts. Employing a difference-in-differences approach we analyze whether consumer assessments (*Zagat* ratings) have responded to *Michelin* quality assessments. While we do not find any significant Michelin-induced increase in perceived food quality, we find strong Michelin effects on service and décor quality. In addition, the inclusion in the Michelin guide induced substantial price increases. While restaurants that were not Michelin-reviewed can raise their prices in response to food quality improvements, service and décor improvement do not payoff. In contrast, Michelin-reviewed restaurant enjoy substantial returns only to service and décor improvement. Our results suggest that expert opinion on the New York City restaurant market exerts a negative externality on gourmets by giving restaurants incentives to invest mainly in service and décor leading to higher prices.

Keywords : Consumer preferences, Expert opinions, Natural experiment, Restaurants.
JEL Codes : D11, L15, L66.

I. INTRODUCTION

In the presence of information asymmetries consumers often rely on expert opinion to guide their purchase decision. An increasing body of economic literature analyzes the effect of critical assessments on prices and quantity consumed for a wide variety of experience goods such as wine, movies, hotel rooms or books. All of these papers analyze the outcome of influenced quality perception of consumers.

Our paper is less focused on the question whether expert opinion impacts quantity or price of the good in question but rather examines consumers' quality perceptions and their possible changes directly. We analyze whether suddenly appearing expert opinion, on a market with long-standing published consumer-assessed quality evaluations, can alter consumers' quality perception and subsequently change prices? Will consumers stick to their original assessments or will they herd towards the expert's opinion?

We investigate this question by referring to restaurants in New York City and exploiting a natural experiment. As the undisputedly leading restaurant guide,[1] *Zagat* has rated New York City's restaurants since 1979. *Zagat* publishes its guidebook once a year by drawing on consumer surveys. It, therefore, reflects local residents' restaurant preferences, which, until 2005, had been only scantly influenced by experts. There had not been any expert guides to New York City restaurant before 2005. Nationwide expert guides such as the *Mobil Travel Guides*, *Fodor* or the *AAA TourBook* series, for various reasons, have never had any mentionable impact on New York City diners (Ferguson, 2008; Davis, 2012). Although the *New York Times* has published weekly reviews and assigned quality ratings to local restaurants since 1963, the number of reviews has hardly exceeded 50 per year – mostly focused on new openings. In comparison, Zagat reviews about 2000 restaurants per year. This and the fact that the reviews are spread over about 50 *New York Times* issues substantially limited its influence and never challenged *Zagat's* position.[2]

In November 2005, however, with the first release of the red *Michelin Guide New York City*, the first one ever for the United States, *Zagat* faced a serious competition. In its first year, *Michelin* reviewed 471 restaurants and sold more than 100,000 copies (Krummert, 2006). In

[1] On average, about 650,000 copies of the New York City guide are sold per year. In addition, *Zagat* reports 384,000 unique visitors to its paid online subscription service for 2008 (Davis, 2012).

[2] For a comprehensive overview of New York City restaurant reviews, their history, focus and impact, see Davis (2012).

contrast to *Zagat*, *Michelin* relies on experts, i.e., five anonymous professional test eaters. According to Ferguson, while Zagat is a plebiscite, Michelin is a tribunal (Ferguson, 2008).

Although the advent of the *Michelin Guide* was excitingly anticipated in New York,[3] when it finally appeared the results were met with surprise, even with dismay. Many of the city's well-regarded restaurants were not awarded a *Michelin* star while others received unexpected honors (e.g., Kurutz, 2005; Fabricant, 2005b; Cuozzo, 2005a). The press detected a bias toward French-owned venues and the *New York Post* even called the Michelin Guide the "idiot's guide" (Cuozzo, 2005b). "After learning that *Babbo* had received [only] one star, Mario Batali[4] said he didn't think New Yorkers would give much credence to the guide. He was not happy with that ranking, the same as for the *Spotted Pig*, of which he is a part-owner. 'They're blowing it,' he said. 'They can't put the *Spotted Pig* on the same level as *Babbo*'" (Fabricant, 2005b).

What credence did New Yorkers give to the *Michelin Guide*? When tackling this question we do not analyze who of the two assessments, consumer or expert ratings, are closer to (unobserved) "true quality."[5] Instead, we analyze whether *Zagat* ratings have responded to *Michelin* quality assessments and employ a difference-in-differences approach for the years 2004, i.e., two years before the first New York City Michelin edition, and 2007, one year after its publication.

We find that Marco Batali's assessment was correct. Despite the media frenzy in 2005, consumers have not adjusted their food quality perception toward the judgment of the *Michelin Guide* experts. However, and despite Michelin's claim to only be guided by the food's quality and not by décor or service, we find a strong and significant *Michelin* effect on consumers' décor quality perception. At a lower significance level, we also find also find a somewhat smaller effect on *service* quality ratings. It is *a priori* unclear whether these effects are based on demand side imaginations or whether the reviewed restaurants have in fact invested in décor and service enhancements. Since the *Michelin* treatment has not influenced

[3] See, e.g., Florence Fabricant in *The New York Times* (Fabricant, 2005a).
[4] Marco Batali, a Rutgers University economics major, is the chef and owner of New York City icon restaurant *Babbo*. He is best known for his Food Network show *Molto Mario* and his role in *Iron Chef America*.
[5] In contrast to national restaurant guides, *Zagat* ratings reflect the vote of the local population and are based on a local reference. Therefore, *Zagat* ratings are not comparable across cities and rather denote a local ranking (see also Berry and Waldfogel, 2010).

consumers' food quality perception we assume the latter, i.e., *Michelin*-reviewed restaurants have in fact improved their décor.

We also find that restaurants are able to pass this non-food-related investment on to consumers. For the mere inclusion into the *Michelin Guide* we find a substantial marginal price effect of approximately 37% from 2004 to 2007. A price increase of this magnitude may well be in excess of service-induced cost increases. Our analysis thus suggests that expert food reviews may provide restaurants with additional profits for non-food investments.

This remainder of this paper is organized as follows. Section II provides a review of the related theoretical and empirical literature. In Section III we present our data and in Section we outline our econometric approach. Section V reports the results and draws conclusions; Section VI summarizes the main findings.

II. LITERATURE

There is an extensive body of literature on producer and consumer behavior in the presence of experience goods, i.e., goods for which quality cannot be ascertained prior to consumption. Beginning with the analyzes of Nelson (1970; 1974) most of the early literature was theoretical in nature and focused on the firm and its scope of quality signaling through advertising, warranties, reputation or pricing (e.g., Schmalensee, 1978; Shapiro, 1983; Wolinsky, 1983; Milgrom and Roberts, 1986; Bagwell and Riordan, 1991; Tirole, 1996). Parallel, an increasing number of empirical papers appeared analyzing firms' signaling from an economic and a marketing perspective for various goods (e.g., Riesz, 1978; Tellis and Wernerfelt, 1987; Curry and Riesz, 1988; Caves and Greene, 1996; Schnabel and Storchmann, 2010). Many of these papers model consumers' quality perception and its formation implicitly. For instance, Bagwell and Riordan (1991) assume that the credibility and scope of signaling quality through pricing declines as consumer become increasingly informed.

In contrast, there is also a growing body of explicit consumer-related literature focusing on the role of peers and experts on consumer preferences. All of these analyses draw on the assumption that the decisions of other consumers or the assessment of experts contain choice-relevant information. The literature on the influence of peers or "social learning" on individual decisions is based on informal approaches in the psychological literature (e.g., Deutsch and Gerard, 1955; Bandura, 1977). For instance, Becker (1991) developed a formal model in which the demand for a good, here a restaurant meal, depends positively on its aggregate quantity demanded, i.e., on peer demand. Banerjee (1992) and Bikhchandani et al. (1992 and 1998) describe localized conformity, fashions and "herd behavior" as the result of informational cascades where the decision of an individual is influenced by the actions of other individuals before him. Since, in these models, the individual is willing to give up his private information and only follows the preceding peers, the peers' actions do not contain any information and the resulting equilibrium may be inefficient.

McFadden and Train (1996) also hypothesize that consumers learn from other consumers but still utilize their private information. They formalize consumer learning about a new good's quality through a rational decision process between learning from own experience or from the experience of their peers.[6] Morris and Shin (2002) show that, when agents have private information, they might overreact to expert opinion. Compared to a welfare-maximizing planner, consumers put too much weight on the public expert's signal and devalue their private information; this may lead to detrimental welfare effects of expert opinion.

On the empirical side, Salganik et al. (2006) created an artificial ''music market'' in which participants downloaded previously unknown songs. They show that, when providing the treatment group of users with information about other users' music ratings, social learning is a strong determinant of a song's success. Moretti (2011) empirically examines the model of social learning for movie sales from 1982 to 2000. He analyzes movie sales over time compared to prior expectations, measured by the number of screens dedicated to a movie in its opening weekend, and finds a reinforcing pattern. When a movie exceeds expectations in its opening week consumers update their expectations leading to further increasing sales etc. Liu (2006) finds similar results for word-of-mouth effects on movie sales by referring to consumers' internet postings. Cai et al. (2009) set up a randomized natural field experiment in which they assess consumer choices of restaurant menu items. If provided with a (made up) list of "last week's top 5 selling dishes," consumers tend to follow their peers' consumption.

[6] Peer or social learning models are related to the earlier literature on technology adoption, where the spreading of new technologies is based on peer imitation (e.g., Griliches, 1957).

In addition to the literature on social learning from peers, there are also numerous papers that confirm the influence of experts on markets. The effect of experts on market outcomes is hard to measure since expert reviews and "true quality" are often closely correlated. Hence, most studies draw on natural (or real) experiments or make statistical inferences to disentangle the two.

For instance, Ginsburgh (2003) reports that experts significantly determine the market success of movies (through Oscars) and, although to a lesser degree, of books (through the *Pulitzer Prize*). Reinstein and Snyder (2005) examine the impact of critical reviews on movie box revenues and also find positive effects of favorable reviews. Ginsburgh and van Ours (2003) analyze the Queen Elizabeth piano competition in Belgium and find that musicians who are successful in the competition will be rewarded by subsequent market success**.** Similarly, experts affect sales prices for paintings at art auctions by publishing pre-sale estimates in auction catalogues (Bauwens and Ginsburgh, 2000).

Hadj Ali, Lecocq and Visser (2008) analyze the effect of critical points awarded by wine writer Robert Parker on the *en primeur* price of Bordeaux wine. While *en primeur* prices are usually set after the wines have been sampled by Parker, this paper exploits the fact that in 2003 Parker postponed his Bordeaux visit and prices were set after before Parrker's review. They find Parker points to have a significant but small effect on the wine price. Dubois and Nauges (2010) also study the effect of Parker points on *en primeur* prices of Bordeaux wines. They employ a structural empirical approach to disentangle the effect of experts' grades and unobserved quality on the wine price and find a significant "Parker effect." Closer related to our research, Gergaud et al. (2007) find a substantial influence of expert ratings, measured by *Guide Michelin* stars, on Paris restaurant menu prices.[7]

In contrast to price analyses, there are only a few papers that examine the impact of expert opinion on quantity consumed. Drawing on a field experiment in wine retail stores, Hilger et al. (2011) show that favorable expert reviews have a positive influence on quantity consumed, independent of quality. On the other hand, wines that obtained below-average ratings exhibit a decrease in demand. Friberg and Grönqvist (2012) analyze the impact of expert opinion on quantity consumed by referring to the Swedish wine market. They find a substantial and long-

[7] Additional examples of expert opinion leadership can be found, e.g., on sport betting markets (Avery and Chevalier, 1999) or on the stock market (Shleifer, 1986).

lasting effect (more than 20 weeks) of positive reviews. In addition, they also find positive demand effects of neutral reviews and do not find any negative effects of unfavorable reviews.

However, consumers' quality perception is not only influenced by own or others' experience but is also responsive to the respective consumption environment. There is plenty of evidence that consumers make contextual inferences (Kamenica, 2008)[8] and are sensitive to the framing of the decision situation (e.g., Tversky and Kahneman, 1981). For instance, North et al. (1999) show that consumers respond to the kind of music played in a wine store. When French music was played, customers bought more than three times as many French wines than German wines. When German music was played the opposite was true. Wansink et al. (2009) report that the quantity of food we eat is only partially determined by what we were planning on consuming. Environmental factors such as package size, plate size and shape, lighting, variety etc. affect our food consumption volume far more than we realize. Plassmann et al. (2008) even show that marketing actions, such as changes in the price of a product, can affect neural representations of experienced pleasantness. In a laboratory experiment, they scanned human subjects using functional Magnetic Resonance Imaging (MRI) while the subjects tasted identical wines that they believed to be different and sold at different prices. Assumingly expensive wines yield increased reports of flavor pleasantness as well as blood-oxygen-level-dependent activity in the medial orbitofrontal cortex, an area that is widely thought to encode for experienced pleasantness during experiential tasks.

However, contextual quality perceptions are not confined to consumers only. In fact, there is ample evidence that experts' quality assessments may be flawed, biased, inefficient or even utterly made up. It has been shown numerous times that mechanical rules outperform expert advice. For example, Krueger and Wu (1998) suggest that mechanical rules may outperform admission committees when judging the success of economics graduate students. Likewise, Bill James' data-driven baseball talent scouting significantly outperforms intuitive experts' judgments (Ayres, 2007). For the Queen Elizabeth Competition mentioned above, Ginsburgh

---

[8] Kamenica (2008) formalizes a model on consumers' contextual inferences in the presence of product lines. For compromise effect ("neither buy the most nor the less expensive item on the list") and overload choice (i.e., an excessive choice set induces a preference for fewer and simpler options) he shows that the context may contain choice-relevant information reconciling these apparent anomalies with standard rational utility maximization. Alternatively, if the context such as expert opinion does not contain any information about the good, it may be deliver a prestige value by itself, akin to the "warm glow" effect demonstrated to be an important component of charitable giving (Karlan and List, 2007).

and van Ours (2003) report that the order and timing of appearance at the competition are good predictors of the final ranking. Judges systematically prefer those players who perform later in the competition. Ashenfelter and Jones (2012) and Ashenfelter (2008) examine the reliability of wine experts and find that experts' ratings are not efficient predictors for mature Bordeaux wine prices because they fail to incorporate all of the publically available information (such as weather).

Another wine-related example of expert failure concerns the *California State Fair Wine Competition*, the oldest and most prestigious wine competition in North America, awarding Gold, Silver and Bronze medals to submitted wines. Unknown to the wine judges, three identical pourings from the same bottle were included into each 20-wine-flight. While the identity of each wine was concealed a reliable wine judge would assess the quality of identical wines similar. However, as reported by Hodgson (2008), more than 80% of the judges scored the same wine more than two medal groups apart (from Gold to "no medal"); the 20% who did not failed to repeat their performance in the following years. In another paper, Hodgson (2009) statistically analyzes medals awarded in 13 national wine competitions. He was puzzled by the fact that a wine can score a Gold in one competition but receive no medal in others. His study suggests that "winning a Gold medal is greatly influenced by chance alone." (Hodgson, 2009, p. 1).

For the restaurant sector, while experts claim to assess food quality only, they cannot refrain from taking into account non-food framing elements, such as the look of the venue or the choice of wines in the cellars (Chossat and Gergaud, 2003). An extreme example of misleading expert advice is reported by Goldstein (2008) and Ashenfelter et al. (2010). The U.S. wine magazine *Wine Spectator* awarded its prestigious "*Award of Excellence* for having one of the most outstanding restaurant wine lists in the world" to the Milan restaurant *Osteria L'Intrepido*, a venue that does not even exists.

Given the significance of learning from peers and experts and the issues that are inherent to both groups this paper tackles the following question. Against the background of well-established and relatively stable peer reviews and quality perceptions, can suddenly appearing expert opinion exert authoritative influence on consumers and change their quality assessments? To answer this question we refer to a natural experiment.

As New York City's leading restaurant guide, *Zagat* has published consumer reviews of restaurants for more than three decades. Consumers rate a restaurant's food, service and décor. Only in 2006 these consumer ratings faced the competition of considerable expert assessments, i.e., the first publication of the New York City *Michelin Guide*. Michelin only rates the food of a restaurant and oftentimes disagrees with consumer preferences. Did consumers change their assessments, as published in the 2007 *Zagat Guide*, after learning from the experts?

III. DATA

We are interested in whether consumers' restaurant quality perceptions, i.e., *Zagat* ratings, have been influenced by the publication of *Michelin* expert opinion in 2006. The dataset we employ covers all New York City restaurants considered in both the 2004 and 2007 *Zagat* Surveys. These years correspond to two years before and one year after the first publication NYC *Michelin*. We draw on 2004 instead of 2005 data to rule out that our results are influenced by possible *Michelin* announcement effects on consumer assessments or on chef efforts.[9]

In the 2004 issue, Zagat published a total of 1,918 restaurant reviews based on the ratings of almost 26,000 reviewers (Zagat Survey, 2003). In the 2007 issue, it rated 2,014 establishments based on reports of 31,604 restaurant-goers (Zagat Survey, 2006). After removing all chain restaurants from this list, we are left with 1518 observations. For each restaurant *Zagat* provides an average consumer-reported price charged for a reference dinner including one drink and tip for each restaurant. It also provides information on the consumer-perceived quality of food, décor and service on a scale ranging from 0 to 30 points separately for each category. In addition, *Zagat* lists some 90 different ethnic cuisine categories[10] that we

[9] The publication of the first New York City *Michelin* guide was announced in February of 2005.
[10] These categories are the following: Afghan, American New, American Regional, American Traditional, Argentinean, Asian, Australian, Austrian, Bakeries, Barbecue, Belgian, Brasserie, Brazilian, Burmese, Cajun/Creole, Californian, Caribbean, Chinese, Coffeehouses/Dessert, Coffee Shops/Diners, Colombian, Continental, Cuban, Delis/Sandwich Shops, Dim Sum, Dominican, Dutch, Eastern European, Eclectic/International, Egyptian, English, Eritrean, Ethiopian, Filipino, Fish 'n' Chips, French, French Bistro, French New, German, Greek, Hamburgers, Health Food, Hot Dogs, Hungarian, Indian, Indonesian, Irish, Israeli, Italian, Jamaican, Japanese, Jewish, Korean, Lebanese, Malaysian, Mediterranean, Mexican/Tex-Mex, Middle Eastern, Moroccan, Noodle Shops, Nuevo Latino, Persian, Peruvian, Pizza, Polish, Portuguese, Puerto Rican, Russian, Sandwich Shop, Scandinavian, Scottish, Seafood, Soups, South African, South American,

bundled into nine broad categories to avoid singletons: Africa, Asia, Central America, Eastern Europe, Middle East, North America, South American, Western Europe, and Other.

Our treatment group consists of the 471 restaurants that were reviewed in the first *Michelin Guide*, 2006 edition (Michelin Travel Publications, 2005). In contrast to *Zagat*, the *Michelin Guide* claims to review the quality of food only; neither décor nor service quality should affect its rating.[11] *Michelin* rates a restaurant's food quality on a scale from zero to three stars.

Table 1 provides the descriptive statistics for food, service, décor and prices for all restaurants in 2004 and in 2007. For 2007, we also report separate numbers for *Michelin*-reviewed and un-reviewed restaurants. From 2004 to 2007, the mean value in each category, including prices, has increased for the restaurants overall. However, Table 1 also shows that *Michelin*-reviewed restaurants attain higher scores and charge higher prices than non-reviewed restaurants, which can simply be a selection result.

[Insert Table 1 here]

Table 2, therefore, reports the descriptive statistics for each group, control and treatment group, before and after the *Michelin* review. Expectedly, the treatment group was rated higher than the control group in each category, i.e., food, service and décor. This is true before as well as after the treatment. In addition, the mean values for each group and category remained virtually unchanged between 2004 and 2007. In contrast, the average price of the restaurants in treatment group grew significantly after the *Michelin* review. In addition, the dispersion, measured by the Coefficient of Variation (CV)[12] within each category is almost undistinguishable between treatment and non-treatment group on the one hand and over time on the other hand. This also applies to the price dispersion of the non-treatment group in 2004 and 2007 and the treatment group in 2004. After the treatment, however, the reviewed restaurants experienced a substantial increase in price dispersion: the CV of prices grew from 34.1% to 52.4%, suggesting a considerable injection of noise caused by published expert

Southern/Soul, South Western, Spanish, Steakhouses, Swiss, Tapas, Tea Service, Thai, Tibetan, Tunisian, Turkish, Ukrainian, Vegetarian, Venezuelan, Vietnamese.

[11] The *New York Times* quotes Jean-Luc Naret, the director of the *Michelin Guides,* " *Michelin* stars refer only to what is on the plate." (Fabricant, 2005a).

[12] We calculate the CV as standard deviation to the mean, $CV = \sigma / \mu$

opinions. Figure 1 shows the corresponding Kernel density functions, which suggest that treatment group prices have stretched out especially at the high end.

In Table 3 we show the percentage growth rates from 2004 to 2007 in each quality category and in prices separately for treatment and non-treatment group. Although these numbers are uncontrolled for effects such as food ethnicity, they still convey a few interesting developments. First, while we find a perceived food quality improvement for non-treated restaurants of 2.43%, the treatment group exhibits a small decline. Second, and despite the lack in food enhancement, the treatment group shows a substantial price increase of 8.52% while there is almost no increase for unreviewed restaurants.

[Insert Table 2 here]

[Insert Figure 1 here]

[Insert Table 3 here]

This overview, however, disregards any influence of variables such as food ethnicity (cuisine categories), restaurant location, operating hours or payment options. In the following section we will thus employ an econometric model to analyze the *Michelin* effect on the three restaurant quality categories food, service and décor, as well as on restaurant meal prices.

IV. ECONOMETRIC METHODOLOGY

Our econometric analysis relies on three difference-in-differences models, one for each category, i.e., food, service and décor, in order to assess whether the mere inclusion in the *Michelin* guide affected consumer quality assessments. We estimate the following equation:

$$Log(Q_{it}) = \beta_0 + \beta_1 \log(Q_{it-1}) + \beta_2 Mich_i + \beta_3 After_t + \beta_4 Mich_i \times After_t + X_{it}\theta + \varepsilon_{it} \qquad (1)$$

where i denotes individual restaurants and t denotes time. Q_t is a measure of quality of food, service or décor, respectively, measured in period t (i.e., 2007); similarly, Q_{t-1} stands for the quality variables in the prior period, i.e., 2004. Introducing the lagged dependent variable accounts for the persistence of quality over time. $Mich_i$ is a dummy variable that takes on the value one if the restaurant was considered in the 2006 *Michelin* guide (first edition) and zero otherwise. *After* is a time dummy that equals one in the period following the introduction of the guide and zero before. $Mich_i \times After_t$ is the interaction term between the two and measures whether Q has changed differently for those who have been introduced in the guide compared to those who have not (control group). It is also known as the difference-in-differences term. X_{it} is a matrix of control variables such as food ethnicity and some characteristics at the restaurant level (accepts credit card, open after 11pm, open on Sundays, limited number of reviews[13]).

It is an implicit assumption in this setup that the treatment, i.e., being considered in the *Michelin* guide, is random and therefore exogenous. Obviously, it is difficult to argue that being considered in the *Michelin* guide is random and independent of the quality of food (or service or décor, respectively) as reported by consumers in the Zagat guide. We, therefore, suspect an endogeneity bias. To remedy this shortcoming we instrument the treatment itself. Given the geographical clustering of *Michelin*-reviewed restaurants, we use the percentage of treated restaurants in the neighborhood as instruments.

The map provided in Figure 2 shows that all *Michelin*-reviewed restaurants are either in one of two geographical clusters in Manhattan or in two less concentrated groups in Queens and Brooklyn.[14] This spatial concentration suggests that the likelihood of being considered in the *Michelin* guide is not independent of a restaurant's geographical location. We exploit this fact and employ a geographical location variable to instrument for being reviewed by the *Michelin* guide.

[Insert Figure 2 here]

---

[13] *Zagat* discloses if a restaurant received only a low number of reviews.

[14] Aside from these clusters, there is only one isolated *Michelin*-reviewed restaurant in Forest Hills, Queens.

In addition to the geographical location of the restaurant, we also explore other possible instruments for the *Michelin* treatment. We examine ZIP-code level data of various demographic and economic data that may serve as appropriate instruments for the treatment variable.[15] In particular, we employ size and racial composition of the population, per capita income, population share under the poverty line, share of full-service restaurants as well as the number of wine and liquor stores per capita.

Since almost all *Michelin*-starred restaurants are located in upscale neighborhoods we expect these variables to be valid instruments for the treatment variable. In other words, we assume a direct relation between the regional concentration of *Michelin* restaurants and their environment (wealth/poverty and interest of the local population for fine wine and food). In addition, since all restaurants in the sample were already established when the *Michelin* guide was introduced, the instruments should be exogenous.

As will be shown later, the statistical tests tend to strongly support our intuition and show that our instruments are neither weak nor endogenous.

Defining neighbors and instruments

In order to define neighbors, we identify the geographical coordinates of all restaurants[16] and compute the distance between all pairs of observations. The smallest maximum distance between two restaurants in the dataset is 19.44 miles, the largest minimum distance is 2.11 miles and the general average distance between two restaurants is 3.64 miles.[17] We attribute proximity spatial weights as follows:

$$w_{ij} = \begin{cases} 0 & \text{if } d_{ij} \notin [l_b, u_b] \\ 1/d_{ij}^{f} & \text{if } d_{ij} \in [l_b, u_b] \end{cases}$$

[15] There are 176 ZIP codes in New York City.

[16] The coordinates are available in decimal degrees from www.maporama.com and are converted into distances (*Km*) to the equator and to the Greenwich meridian using the formula: $distance = \frac{6378.137 \cdot \pi \cdot degrees}{180}$

[17] The maximum distance between two restaurants is little informative since it merely reports the spatial spread of restaurants in New York City. Similarly, the minimum distance is virtually zero for adjacent restaurants. The largest minimum distance gives us an idea of the minimal radius needed for all restaurants to have at least one neighbor. The smallest maximum distance, on the other hand, reflects the spatial spread of restaurants compared to the central restaurant.

where (i,j) denotes a pair of locations, d_{ij} stands for the Euclidean distance between restaurant i and j, l_b and u_b denote the lower and upper bound of the specified distance band, respectively, and f is a positive friction parameter that is set exogenously. The friction parameter determines the rate of devaluation for neighbors compared to the geographic distance. A parameter value of one denotes that the importance of the neighborhood effect is linearly decaying in distance. A friction parameter larger than one suggests that neighborhood effects decline faster than the geographic distance and vice versa.

Since in New York City, the monetary transportation cost is virtually independent of distance traveled while time spent depends on distance, we set the friction parameter equal to 0.8 suggesting below-proportional neighbor depreciations compared to the geographic distance. However, our empirical results are not overly sensitive to different parameter values.

Finally, the values in the weighting matrix are standardized in order to ensure that the sum of all elements per row equals one. A restaurant i is considered a neighbor of restaurant j if the distance between i and j does not exceed 10 km (i.e. l_b=0 and u_b=10).[18]

We can now easily calculate the average number of *Michelin* restaurants in the neighborhood of each restaurant (weighted by the distance) by multiplying the weighting matrix (W) by the vector identifying the *Michelin* restaurants. In other words, the frequency of "*Michelin* restaurants" in the neighborhood of each restaurant is defined by $W \cdot Mich$ (vector *WMich*). This variable is the first instrument we use for the treatment. The second instrument we consider is a dummy variable that is equal to one if only a small number of customers reviewed the restaurant (*Low2004*); this variable is provided by the Zagat guide. We hypothesize that *Michelin* can afford to disregard unknown restaurants. However, restaurants with a large number of customer reviews may enjoy an increased likelihood of being selected in the guide. Since *Zagat* refers to the number of 2004 reviews, i.e., well before the announcement of the *Michelin* launch, we deem this variable exogenous.

To summarize, the endogenous right hand side variables are *Mich* and *(Mich x After)*. The available instruments are *WMich*, *WMich* interacted with *After* (which is exogenous) and *Low2004*. Since we employ more instruments than we have endogenous variables we test for

[18] The distance of 10 km was selected to ensure that each restaurant has at least one neighbor.

their redundancy, over-identification (i.e., exogeneity), and weakness. We perform these tests for each model, i.e., for food, décor and service, by drawing on the Hansen J-statistic (exogeneity test), the Kleibergen-Paap rk LM statistic (relevance test) and the Kleibergen-Paap rk Wald F statistic (weakness test).

V. RESULTS

A. Impact on Quality

In Table 4 we report the results of the model described in equation (1) with respect to the quality of food, service and decor. The results of the models using instruments are given in columns (4) to (6). For comparison reasons we also report the result of the simple OLS equations in columns (1) to (3). Note that the treatment variable is specified as a 0-1 dummy variable, i.e., we only distinguish between *reviewed* and *not reviewed*.

[Insert Table 4 here]

As already suggested by Table 1 and 2 and Figure 1, the quality of all three variables increased for all restaurants (see *After*) from 2004 to 2007. Unexpectedly, however, we do not find any significant evidence for the assumption that *Michelin*-reviewed restaurants are of superior quality (see *Michelin*). Although the OLS estimates for the food variable are significant at the 5%-level, all IV estimates are insignificant.

Even more surprising are the estimates for the interaction term (see *Michelin x After*), which denotes the treatment effect. While the OLS models report significant quality improvements for all three variables, the confidence levels decline substantially when referring to the 2SLS models. Only the décor (5%-level) and service (10%-level) variables remain significant. The food variable does not exhibit any *Michelin* treatment-induced quality increase. Although the interaction term is significantly different from zero for all three variables, the largest effect is on décor and service. In general, improved perceived quality can be the result of supply side investments or merely imagined by consumers (demand side). Since we find only *Michelin*-induced perceived service and décor quality improvements but no effects on food quality, we assume that *Michelin*-reviewed restaurant in fact invested in their service and décor.

In Table 5, we report the results of various tests for exogeneity, relevance and weakness of our instruments. Note that we chose different combinations of instruments depending on the resulting test statistics and that we instrument both *Michelin* and *Michelin x After*. The variable "limited number of reviews" refers to a restaurant's (quantitative) unpopularity. The geography variable, as described above, denotes the regional concentration of reviewed restaurants; the number of wine stores, the share of population below the poverty line and the share of full service restaurants are by ZIP code and reflect various aspects of neighborhood desirability.

[Insert Table 5 here]

For the food variable, we calculate the Hansen J-statistic to check for the exogeneity of the instruments. The resulting value of 5.17 is below the critical χ^2 value for three degrees of freedom (7.815). We hence do not reject the null hypothesis that the instruments are exogenous. To check for the relevance of the instruments, we rely on the Kleibergen-Paap rk LM statistic, which equals 39.21 for the food model. This value is well above the critical χ^2 value for four degrees of freedom, which is 9.488. Therefore, we reject the null that the model is underidentified. Finally, to test for the weakness of the set of instruments, we compute the Kleibergen-Paap rk Wald F statistic. The resulting value of 6.797 for the food model is lower (larger) than the critical value of 8.78 (5.91) tabulated by Stock and Yogo (2005) for a 20% (10%) maximal IV relative bias. We find similar results for the service and decor equations.

Table 6 and 7 report the results and instrument statistics when we replace the *Michelin* treatment dummy variable with an ordered variable that takes on the value 0 for not reviewed, 1 for reviewed but no star, 2 for one star, 3 for two stars and 4 for three stars. The ordered variable thus postulates a constant marginal effect of each additional *Michelin* star on the various quality variables. Overall, the results are very similar to the findings when using a treatment dummy variable.

[Insert Table 6 here]

[Insert Table 7 here]

Although our results that an expert guide -- claiming to rate food quality only -- changes consumers' perception of décor and service but not food quality assessments are somewhat surprising, they are not entirely new. Chossat and Gergaud (2003) and Gergaud et al. (2007) show that *Michelin* evaluations in France are not solely driven by food quality but also influenced by non-food characteristics such as décor and service. Johnson and Surlemot (2005) interviewed chef-owners of *Michelin* starred restaurants in France, Belgium, Switzerland and the UK and report that receiving a *Michelin* star places enormous pressure on the owner. Massive efforts and investments are due in order to retain the recently gained (first, second or third) *Michelin* star. Since these investments include service and décor it seems to be commonly understood among restaurateurs that *Michelin* ratings – in contrast to their claim - are influenced by service and décor. These findings suggest that the higher service and décor quality may not be imagined by consumers but may rather be the result of the owner's effort.

Investments in service and décor are expensive and may only be justified if they yield higher revenue. There is some anecdotal evidence that *Michelin* stars demand a premium and are thus worth being retained. Eric Ripert, chef and owner of *Le Bernardin*, one of only three New York City restaurants that received three *Michelin* stars in 2006, reports revenue increases of at least 15% (Davis, 2012). Johnson and Surlemot (2005) find similar values for European *Michelin* starred venues. In an analysis of French *Michelin* reviewed restaurants from 1970 to 1994 Snyder and Cotter (1998) find a close relationship between investments, especially in ambience, *Michelin* stars and prices. In particular, the loss of a *Michelin* star is often predated by receding investments and lower prices.

B. Impact on Prices

In Table 8 we show the impact of the *Michelin* treatment (i.e., being included in the Guide) on menu prices. The model specification is identical with the one for the *Zagat* quality assessments (see equation 2); we only substituted the logarithm of menu prices for the *Zagat* variable as dependent variable. Exogeneity, under identification and weak identification tests for the selected instruments are reported in Table 9. Accordingly, when employing a simple 1-0 dummy variable for the inclusion in the *Michelin Guide* we find treatment-induced price increases of approximately 37%. When using an ordered 0-1-2-3-4 treatment variable we find a marginal effect of 0.22. That is, one *Michelin* star yields an approximate price premium of 44% while three stars will cause price increases of about 88%.

[Insert Table 8 here]

[Insert Table 9 here]

We are further interested in examining whether these price increases are related to food, service or décor quality improvements and whether there is a difference between *Michelin*-reviewed and un-reviewed restaurants. We select all unique restaurants for which we have price and quality data for 2004 and 2007 and regress the nominal price difference on the respective quality difference and a constant term. In this fashion we ran 12 different regressions; Table 10 displays the results. In the Columns denoted “All” we draw on 700 un-reviewed and 338 reviewed restaurants. For the group of un-reviewed restaurants we find positive marginal effects of food, service and décor quality changes. However, only décor changes exert a modestly significant effect on price changes; the other quality variables remain insignificant. In contrast, the prices of *Michelin*-reviewed respond to service and décor changes at the 1% and 2% significance level, respectively. In addition, the marginal effects for *Michelin*-treated restaurants are between three times (décor) and 8.5 times (service) larger than for un-treated restaurants. These effects are remarkable given that the average price difference between treated an untreated restaurant was only 38% in 2004 (see Table 1). In contrast, food quality improvements do not seem to trigger any price increases.

[Insert Table 10 here]

These results suggest that *Michelin*-treated restaurants improve the quality of their service and décor but not the quality of their food in order to raise prices. On the other hand, prices of untreated restaurants are only weakly driven by décor improvements.

However, when regressing price changes only on changes in one quality dimension, we disregard possible changes in the other quality variables and may just confound the respective marginal effects. We, therefore, restrict our sample to only these restaurants that exhibited a change in only one quality variable. For instance, when regressing price changes on food quality changes, we only refer to restaurants for which décor and service has not changed by more than one Zagat quality point. While this procedure allows us to isolate the respective quality effects on prices, our sample size now drops to between 439 and 473 for untreated and between 239 and 252 for *Michelin*-treated restaurants. The corresponding results, as reported in the Columns denoted “Restricted Sample” in Table 10, confirm our prior findings, although at a somewhat lower significance level. On the one hand, price changes of restaurants that have not been reviewed by the *Michelin* guide are determined by food quality changes; service and décor improvement are irrelevant. On the other hand, menu price increases of *Michelin*-reviewed restaurants are fully determined by service and décor quality and completely detached from food quality changes.

Figure 3 shows the scatterplots and local polynomial smooths for the restricted data samples, i.e., when keeping the other quality variables virtually constant. The three left-hand columns report the price-quality change relationships for *Michelin* restaurants. While even 3 Zagat-point food improvements have not yielded any price increases, the returns to larger service improvements have been substantial. For un-reviewed restaurants, reported in the three columns on the right-hand, the opposite is true. While food quality improvements have translated into higher menu prices, even significant service and décor improvements have not paid off.

[Insert Figure 3 here]

Against this background, one needs to keep in mind that the *Michelin* guide claims to only assess a restaurant's food quality and disregards its service and décor. Our results suggest that food expert reviews initiate service and décor improvements leading to higher prices and thus exerting negative external effects for gourmets.

VII. SUMMARY AND CONCLUSIONS

In this paper we analyze whether consumers' quality perception is influenced by newly appearing expert opinion. We investigate this question by referring to restaurants in New York City and exploiting a natural experiment. As the leading restaurant guide *Zagat* has rated New York City's restaurants since 1979 by surveying more than 30,000 restaurant goers per year. In 2005, with the first release of the red *Michelin Guide New York City*, *Zagat* faced a serious competition. In contrast to *Zagat*, *Michelin* relies on expert eaters. Employing a difference-in-differences approach we analyze whether consumer assessments (*Zagat* ratings) have responded to *Michelin* quality assessments. While we surprisingly do not find any significant *Michelin*-induced increase in perceived food quality, being *Michelin*-reviewed exerts significantly positive effects on service and décor quality. We assume that the service and décor effects are not imagined but rather based on real restaurant investments. This conjecture is in line with the idea that improving food quality usually takes more time and is a more complicated (because artistic) task than improving décor that just requires instant investments in renovations and/or improving service hiring more and/or better waiters.

We find that the inclusion in the *Michelin* guide induced substantial price increases. While restaurants that were not *Michelin*-reviewed can raise their prices in response to food quality improvements, service and décor improvement do not payoff. In contrast, *Michelin*-reviewed restaurant, on the other hand, enjoy substantial returns only to service and décor improvement. Our results suggest that expert opinion on the New York City restaurant market exerts a negative externality on gourmet by giving restaurants incentives to invest in service and décor leading to higher prices for the same food.

REFERENCES

Ashenfelter, O. (2008). Predicting the quality and prices of Bordeaux wines. *The Economic Journal*, 118(529), F174-F184.

Ashenfelter, O., Goldstein, R. and Riddell, C (2010). Do expert ratings measure quality? The case of restaurant wine lists. Presentation at the 4th Annual Conference of the American Association of Wine Economists at UC Davis. Mimeo.

Ashenfelter, O. and Jones, G. (2012). The demand for expert opinion: Bordeaux wine. *Journal of Wine Economics, 7(2), forthcoming*.

Avery, C. and Chevalier, J. (1999). Identifying investor sentiment from price paths: the case of football betting. *Journal of Business*, 72(4), 494-521.

Ayres, I. (2007). *Super Crunchers. Why Thinking-by-Numbers Is the New Way to Be Smart.* New York: Bantam Books.

Bagwell, K. and Riordan, M.H. (1991). High and declining prices signal product quality. *American Economic Review*, 81, 224-239.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice Hall.

Banerjee, A. (1992). A simple model of herd behavior. *Quarterly Journal of Economics*, 107(3), 797-817.

Bauwens, L. and Ginsburgh, V. (2000). Art experts and auctions. Are pre-sale estimates unbiased and fully informative? *Louvain Economic Review*, 66, 131–144.

Berry, S. and Waldfogel, J. (2010). Product quality and market size. *Journal of Industrial Economics*, 58(1), 1-31.

Becker, G.S. (1991). A note on restaurant pricing and other examples of social influences on price. *Journal of Political Economy*, 99(5), 1109-1116.

Bikhchandani, S, Hirshleifer, D. and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5), 992-1026.

Bikhchandani, S., Hirshleifer, D. and Welch, I. (1998). Learning from the behavior of others: conformity, fads, and informational Cascades. *Journal of Economic Perspectives*, 12(3), 151-170.

Cai, H., Chen, Y. and Fang, H. (2009). Observational learning: evidence from a randomized natural field experiment. *American Economic Review*, 99(3), 864-882.

Caves, R.E. and Greene, D.P. (1996). Brands' quality levels, prices, and advertising outlays: empirical evidence on signals and information cost. *International Journal of Industrial Organization*, 14, 29-52.

Chossat, V. and Gergaud, O. (2003). Expert opinion and gastronomy: the recipe for success. *Journal of Cultural Economics*, 27, 127–141.

Cuozzo, S. (2005a). A pinch of bias leaves bitter taste. *New York Post*, November 2.

Cuozzo, S. (2005b). The idiot's guide; Michelin diner book bites. *New York Post*, November 7.

Curry, D.J. and Riesz, P.C. (1988). Prices and price/quality relationships: a longitudinal analysis. *Journal of Marketing*, 52, 36-51.

Davis, M. (2012). A taste of New York: restaurant reviews, food discourse, and the field of gastronomy in America. (PhD Thesis, New York University, 2009). American Association of Wine Economists AAWE, *AAWE Working Paper* No. 106.

Deutsch, M. and Gerard, H.B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, 51, 629-636.

Dubois, P. and Nauges, C. (2010). Identifying the effect of unobserved quality and expert reviews in the pricing of experience goods: empirical application on Bordeaux wine. *International Journal of Industrial Organization*, 28, 205-121.

Fabricant, F. (2005a). The tire man eats New York. *New York Times*, February 23.

Fabricant, F. (2005b). Is New York worth a trip? Oui. *New York Times*, November 2.

Ferguson, P.P. (2008). Michelin in America. *Gastronomica*, 8(1), 49-55.

Friberg, R. and Grönqvist, E. (2012). Do expert reviews affect the demand for wine. *American Economic Journal: Applied Economics*, 4(1), 193–211.

Gergaud, O., Montano, L. and Verardi V. (2007). Stardust over Paris gastronomic restaurants. *Journal of Wine Economics*, 2(1), 24-39.

Ginsburgh, V. (2003). Awards, success and aesthetic quality in the arts. *Journal of Economic Perspectives*, 17(2), 99-111.

Ginsburgh, V. and van Ours, J. (2003). Expert opinion and compensation: evidence from a musical competition. *American Economic Review*, 93, 289-296.

Goldstein, R. (2008). What does it take to get a Wine Spectator Award of Excellence? *Blind Taste, a critical review of food, drinks, culture, and cognition. Blog. Online at* http://blindtaste.com/2008/08/15/what-does-it-take-to-get-a-wine-spectator-award-of-excellence/

Griliches, Z. (1957). Hybrid corn: an exploration in the economics of technical change. *Econometrica*, 25(4): 501-522.

Hadj Ali, H., Lecocq, S. and Visser, M. (2008). The impact of gurus: Parker grades and *en primeur* wine prices. *Economic Journal*, 118, F158-F173.

Hilger, J., Rafert, G. and Villas-Boas, S. (2011). Expert opinion and the demand for experience goods: an experimental approach in the retail wine market. *Review of Economics and Statistics***,** 93(4), 1289-1296.

Hodgson, R.T. (2008). An examination of judge reliability at a major U.S. wine competition. *Journal of Wine Economics*, 3(2), 105-113.

Hodgson, R.T. (2009). An analysis of the concordance among 13 U.S. wine competitions. *Journal of Wine Economics*, 4(1), 1-9.

Johnson, C. and Surlemont, B. (2005). Behind the stars : a concise typology of Michelin restaurants in Europe. Cornell Hotel and Restaurant Administration Quarterly, 46(2), 170-187.

Kamenica, E. (2008). Contextual inference in markets: on the informational content of product lines. *American Economic Review*, 98(5), 2127–2149.

Karlan, D. and List, J.A. (2007). Does price matter in charitable giving? Evidence from a large-scale natural field experiment. *American Economic Review*, 97(5), 1774-1793.

Krueger, A.B. and Wu, S. (1998). Forecasting successful economics graduate students. Princeton University Industrial Relations Section Working Paper No. 403.

Krummert, B. (2006). Making a list, and checking it twice. *Restaurant Hospitality*, 30, (October 31, 2006).

Kurutz, S. (2005). Neighborhood report: Union Square; she’s a belle of the city, but the French are blasé. *New York Times*, November 13.

Liu, Y. (2006). Word of mouth for movies: its dynamics and impact on box office revenues. *Journal of Marketing*, 70(3), 74-89.

McFadden, D.L. and Train, K. (1996). Learning from self and others. *Journal of Political Economy*, 104(4), 684-703.

Michelin Travel Publications (2005). New York City 2006 : selection of restaurants and hotels. Greenville, SC : Michelin Travel Publications

Milgrom, P. and Roberts, J. (1986). Price and advertising signals of product quality. *Journal of Political Economy*, 94, 796-821.

Moretti, E. (2011). Social learning and peer effects in consumption: evidence from movie sales. *The Review of Economic Studies*, 78(1), 356-393.

Morris, S. and Shin, H.S. (2002). Social value of public information. *American Economic Review*, 92(5), 1521-1534.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78, 311-329.

Nelson, P. (1974). Advertising as information. *Journal of Political Economy*, 81, 729-754.

North, A., Hargreaves, D. and McKendrick, J. (1999). The influence of in-store music on wine selections. *Journal of Applied Psychology*, 84(2), 271-276.

Plassmann, H., O’Doherty, J., Shiv, B. and Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences PNAS*, 105(3), 1050-1054.

Reinstein, D.A. and Snyder, C.M. (2005). The influence of expert reviews on consumer demand for experience goods: a case study of movie critics. *Journal of Industrial Economics*, 53(1), 27-51.

Riesz, P.C. (1978). Price versus quality in the market place. *Journal of Retailing*, 58, 15-28.

Salganik M.J., Dodds, P.S. and Watts, D.J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311, 854-856.

Schmalensee, R. (1978). A model of advertising and product quality. *Journal of Political Economy*, 86, 485-503.

Schnabel, H. and Storchmann, K. (2010). Prices as quality signals: evidence from the wine market. *Journal of Agricultural & Food Industrial Organization*, 8, Article 2.

Shapiro, C. (1983). Premiums for high quality products as returns to reputation. *The Quarterly Journal of Economics*, 98, 659-680.

Shleifer, A. (1986). Do demand curves for stocks slope down? *Journal of Finance*, 41, 579–590.

Snyder, W. and Cotter, M. (1998). The Michelin Guide and restaurant pricing strategies. *Journal of Restaurant & Foodservice Marketing*, 3(1), 51-67.

Stock, J.H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D.W.K. Andrews and J.H. Stock, eds. Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge: Cambridge University Press, 80–108.

Tellis, G.J. and Wernerfelt, B. (1987). Competitive price and quality under asymmetric information. *Marketing Science*, 6, 240-253.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies*, 63, 1-22.

Tversky, A. and Kahneman, D (1981). The framing of decisions and the psychology of choice. *Science*, 30(211), 453-458.

Wansink, B., Just, D.R. and Payne, C.R. (2009). Mindless eating and healthy heuristics for the irrational. *American Economic Review, Papers & Proceedings*, 99(2), 165-169.

Wolinsky, A. (1983). Prices as signals of product quality. *Review of Economic Studies*, 50, 647-658.

Zagat Survey (2003). New York City Restaurants 2004. New York: Zagat Survey.

Zagat Survey (2006). New York City Restaurants 2007. New York: Zagat Survey.

Table 1

Descriptive Statistics for Zagat Food, Service, Décor Ratings and Price

2004 and 2007, by Reviewed and Unreviewed

	Number of observations	Mean	Std. Deviation	Minimum	Maximum
			Food		
2004 all	1497	20.46	2.82	9	28
2007 all	1516	20.97	2.58	10	29
2007 unreviewed	1047	20.28	2.34	10	29
2007 reviewed	469	22.52	2.43	16	28
			Service		
2004 all	1497	18.20	3.17	8	30
2007 all	1518	18.47	3.08	8	29
2007 unreviewed	1049	17.75	2.84	8	29
2007 reviewed	469	20.07	2.99	10	29
			Décor		
2004 all	1494	16.64	4.56	2	28
2007 all	1517	16.79	4.50	3	29
2007 unreviewed	1049	15.74	4.28	3	29
2007 Michelin reviewed	468	19.14	4.07	5	28
			Price		
2004 all	1497	38.14	14.94	5	185
2007 all	1515	40.69	20.64	5	446
2007 unreviewed	1048	35.60	14.52	5	215
2007 Michelin reviewed	467	52.12	26.85	14	446

Source: Zagat Survey (2003 and 2006).

Table 2

Descriptive Statistics Food, Service and Décor Quality and Price

Treatment and Non-Treatment Group 2004 and 2007

	Number of observations	Mean	Std. Dev. (CV in %)	Minimum	Maximum
			Food		
Non-Treatment 2004	1109	19.72	2.55 (12.9)	9	28
Non-Treatment 2007	1002	20.20	2.34 (11.6)	10	29
Treatment 2004	388	22.59	2.44 (10.8)	15	28
Treatment 2007	514	22.47	2.38 (10.6)	16	28
			Service		
Non-Treatment 2004	1109	17.53	2.96 (16.9)	8	30
Non-Treatment 2007	1003	17.60	2.78 (15.8)	8	29
Treatment 2004	388	20.12	2.98 (14.8)	10	27
Treatment 2007	515	20.14	2.96 (14.7)	10	29
			Décor		
Non-Treatment 2004	1106	15.84	4.38 (27.7)	2	28
Non-Treatment 2007	1003	15.62	4.24 (27.1)	3	29
Treatment 2004	338	18.94	4.28 (22.6)	5	28
Treatment 2007	514	19.07	4.11 (21.6)	4	28
			Price		
Non-Treatment 2004	1109	34.70	12.73 (36.7)	5	93
Non-Treatment 2007	1002	34.86	12.77 (36.6)	5	85
Treatment 2004	388	47.99	16.38 (34.1)	16	185
Treatment 2007	513	52.08	27.28 (52.4)	14	446

Source: Zagat Restaurant Guide New York City, 2004 and 2007.

Table 3
Percentage Change in Quality and Price from 2004 to 2007
Treatment and Non-Treatment Group

		Change
Food	Non-Treatment	2.43%
	Treatment	-0.53%
Service	Non-Treatment	0.40%
	Treatment	0.10%
Décor	Non-Treatment	-1.39%
	Treatment	0.69%
Price	Non-Treatment	0.46%
	Treatment	8.52%

Table 4

Perceived Quality Difference-in-Difference Equations

Dependent variable: Log of perceived quality; Treatment : Michelin 2006 (dummy)

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
	Food	Décor	Service	Food	Décor	Service
Michelin	0.00**	0.00	0.00	0.01	-0.02	-0.00
	(2.15)	(0.51)	(0.57)	(1.43)	(-0.57)	(-0.42)
Michelin x After	0.10***	0.18***	0.12***	0.04	0.29**	0.11*
	(16.20)	(12.76)	(13.64)	(0.79)	(2.48)	(1.65)
After	2.88***	2.44***	2.58***	2.86***	2.50***	2.62***
	(147.46)	(60.66)	(82.16)	(66.66)	(52.04)	(62.37)
Lagged Dependent (2004)	0.96***	0.90***	0.90***	0.94***	0.92***	0.91***
	(152.54)	(66.46)	(84.01)	(65.36)	(51.79)	(54.41)
Open after 11 PM	-0.02***	-0.00	-0.03***	-0.02***	0.01	-0.03***
	(-5.40)	(-0.10)	(-5.41)	(-5.40)	(1.20)	(-4.69)
No Credit Cards Accepted	0.02***	-0.16***	-0.06***	0.02**	-0.16***	-0.07***
	(4.45)	(-7.89)	(-6.87)	(2.47)	(-6.35)	(-5.59)
Closed on Sunday	-0.03***	-0.03***	-0.04***	-0.03***	-0.02	-0.04***
	(-6.11)	(-2.73)	(-6.25)	(-4.70)	(-1.27)	(-5.22)
Limited # of reviews (2004)	-0.00**	-0.00	-0.00			
	(-2.51)	(-0.35)	(-0.46)			
Constant	0.17***	0.28***	0.34***	0.22***	0.21***	0.31***
	(7.03)	(6.10)	(9.44)	(4.58)	(3.60)	(5.72)
Food ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3012	3010	3014	2475	2473	2498
R-squared	0.65	0.61	0.57	0.64	0.60	0.56

Dependent variable: ln of perceived quality; Treatment: Not in Michelin = 0 ; In Michelin = 1.
Robust z-statistics in parentheses; z-statistics are based on restaurant-clustered standard errors;
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$;

Table 5

Exogeneity, Under Identification and Weak Identification Tests

	Food (4)	Décor (5)	Service (6)
Under identification test :			
Kleibergen-Paap rk (LM statistic)	39.205	32.547	35.508
	(0.000)	(0.000)	(0.000)
Weak identification test :			
Kleibergen-Paap stat.	6.797	6.523	8.199
IV Relative bias	10-20%	10-20%	5-10%
Exogeneity of instruments :			
Hansen's overidentification test	5.169	2.649	4.675
	(0.160)	(0.266)	(0.097)
Instruments:			
Limited reviews	Yes	Yes	Yes
Geography	Yes	Yes	Yes
Wine stores	Yes	No	No
Poverty share	Yes	Yes	No
Proportion of full service rest.	No	No	Yes

Table 6

Perceived Quality Equations (Difference-in-differences) for Ordered Treatment

Dependent variable: Log of perceived quality; Treatment : Michelin 2006 (ordered)

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
	Food	Décor	Service	Food	Décor	Service
Michelin	0.00**	0.00	0.00	0.01	-0.01	-0.00
	(2.37)	(0.55)	(0.58)	(1.41)	(-0.47)	(-0.42)
Michelin x After	0.09***	0.16***	0.11***	0.02	0.24**	0.09*
	(19.52)	(14.83)	(16.64)	(0.66)	(2.49)	(1.65)
After	2.88***	2.45***	2.58***	2.86***	2.50***	2.62***
	(151.28)	(61.15)	(84.11)	(67.42)	(52.85)	(62.51)
Lagged Dependent (2004)	0.96***	0.90***	0.90***	0.94***	0.92***	0.91***
	(156.20)	(66.98)	(86.02)	(65.47)	(51.68)	(53.95)
Open after 11 PM	-0.02***	-0.00	-0.03***	-0.02***	0.01	-0.03***
	(-5.45)	(-0.10)	(-5.41)	(-5.62)	(1.15)	(-4.89)
No Credit Cards Accepted	0.02***	-0.16***	-0.06***	0.02**	-0.16***	-0.07***
	(4.49)	(-7.91)	(-6.87)	(2.45)	(-6.51)	(-5.74)
Closed on Sunday	-0.02***	-0.03**	-0.04***	-0.02***	-0.01	-0.04***
Limited # of reviews (2004)	-0.00**	-0.00	-0.00	-	-	-
	(-2.40)	(-0.32)	(-0.48)	-	-	-
Constant	0.17***	0.28***	0.33***	0.22***	0.21***	0.31***
	(7.06)	(6.02)	(9.34)	(4.57)	(3.59)	(5.66)
Food ethnicity fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3012	3010	3014	2475	2473	2498
R-squared	0.66	0.61	0.58	0.64	0.61	0.57

Robust z-statistics in parentheses; z-statistics are based on restaurant-clustered standard errors; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Treatment : 0 = not in Michelin ; 1 = in Michelin ; 2 = one star ; 3 = two stars ; 4 = three stars.

Table 7

Exogeneity, Under Identification and Weak Identification Tests

	Food (4)	Décor (5)	Service (6)
Under identification test :			
Kleibergen-Paap rk (LM statistic)	38.035	32.688	35.025
	(0.000)	(0.000)	(0.000)
Weak identification test :			
Kleibergen-Paap stat.	7.191	6.871	8.549
IV Relative bias	10-20%	10-20%	5-10%
Exogeneity of instruments :			
Hansen's overidentification test	5.584	2.489	4.757
	(0.134)	(0.288)	(0.093)
Instruments:			
Limited reviews	Yes	Yes	Yes
Geography	Yes	Yes	Yes
Wine stores	Yes	No	No
Poverty share	Yes	Yes	No
Proportion of full service rest.	No	No	Yes

Table 8

Difference-in-Differences Price Equations

Treatment: Michelin 2006 (dummy and ordered)

	Treatment		
	Dummy	**Ordered**	
Michelin	0.52***	0.51***	
	(4.43)	(4.43)	
Michelin × After	0.37**	0.22	
	(2.25)	(1.46)	
After	-.06	-0.05	
	(-1.33)	(-1.03)	
Open after 11 PM	0.02	0.01	
	(0.74)	(0.69)	
No Credit Cards Accept	-0.37***	-0.38***	
	(9.99)	(-10.54)	
Closed on Sunday	-0.16***	-0.15**	
	(5.52)	(-4.71)	
Constant	3.60***	3.61***	
	(51.85)	(54.92)	
Food origin dummies	Yes	Yes	
Observations	2476	2476	
R-squared	0.12	0.17	

Dependent variable: ln of price; Robust z-statistics in parentheses; z-statistics are based on restaurant-clustered standard errors; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; *** $p < 0.01$, ** $p < 0.05$;
Treatment: 0 = not in Michelin; 1 = in Michelin, no star; 2 = one star; 3 = two stars; 4 = three stars.

Table 9

Exogeneity, Under Identification and Weak Identification Tests

(Price equations)

	Dummy	Ordered
Under identification test :		
Kleibergen-Paap rk (LM statistic)	38.020	34.123
	(0.000)	(0.000)
Weak identification test :		
Kleibergen-Paap stat.	7.927	7.740
IV Relative bias	5-10%	5-10%
Exogeneity of instruments :		
Hansen's overidentification test	2.010	2.826
	(0.366)	(0.243)
Instruments:		
Limited reviews	Yes	Yes
Geography	Yes	Yes
Wine stores	Yes	No
Poverty share	No	No
Proportion of full service rest.	No	No

Table 10
Determinants of Price Changes
Dependent Variable: Nominal Price Change from 2004 to 2007

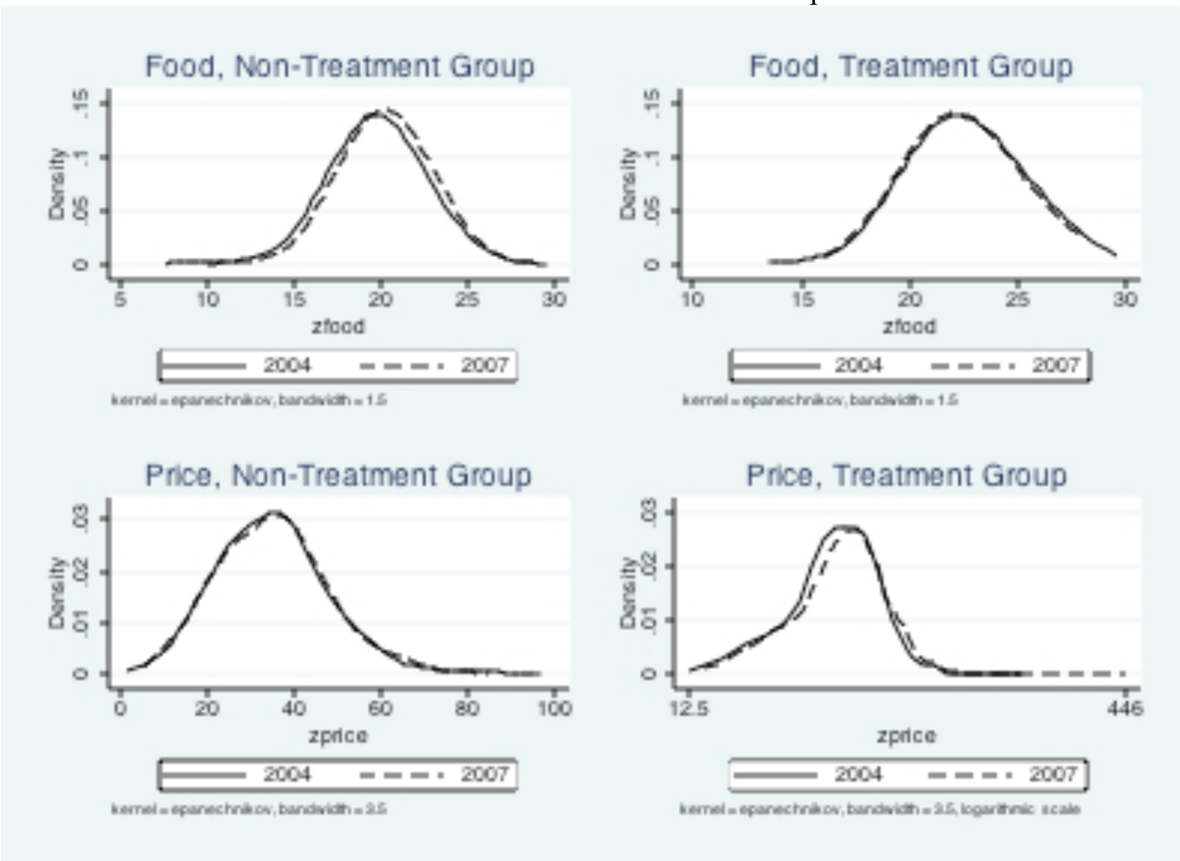
	Not Michelin Reviewed		Michelin Reviewed	
	All	Restricted Sample[a]	All	Restricted Sample[a]
Change in Food Quality	0.0139 (700; 0.18)	0.1966+ (439; 1.85)	0.2301 (338; 1.01)	-0.2310 (239; -0.88)
Change in Décor Quality	0.1207+ (700; 1.75)	0.0143 (473; 0.19)	0.3612** (338; 2.45)	0.3015+ (252; 1.90)
Change in Service Quality	0.0491 (700; 0.57)	0.1415 (451; 1.41)	0.4189*** (338; 2.58)	0.3645+ (240; 1.72)

All equations contain a constant term (not reported here); number of observations and robust t-statistics in parentheses (n; t-stats). Significance + (6%), * (5%), ** (2%), *** (1%). [a]) Sample is restricted to observations with quality changes of one or less Zagat point for the other quality variables. For instance, the impact of food quality changes is measured for restaurants that have not changed the service and decor quality by more than one point up or down.

Figure 1

Kernel Density Function for Food and Price 2004 and 2007

Treatment and Non-Treatment Group



`

Figure 2

Michelin Restaurants in New York City

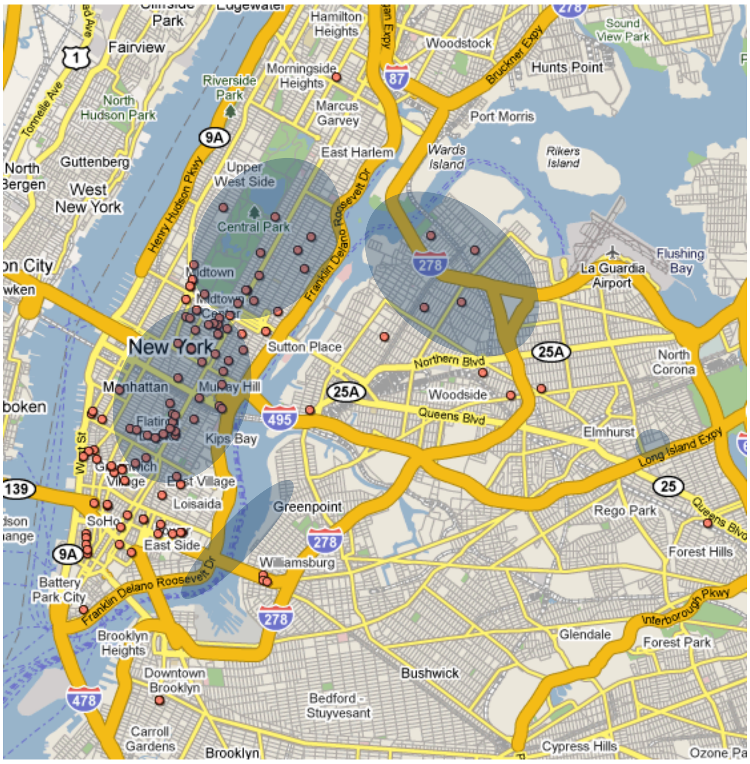
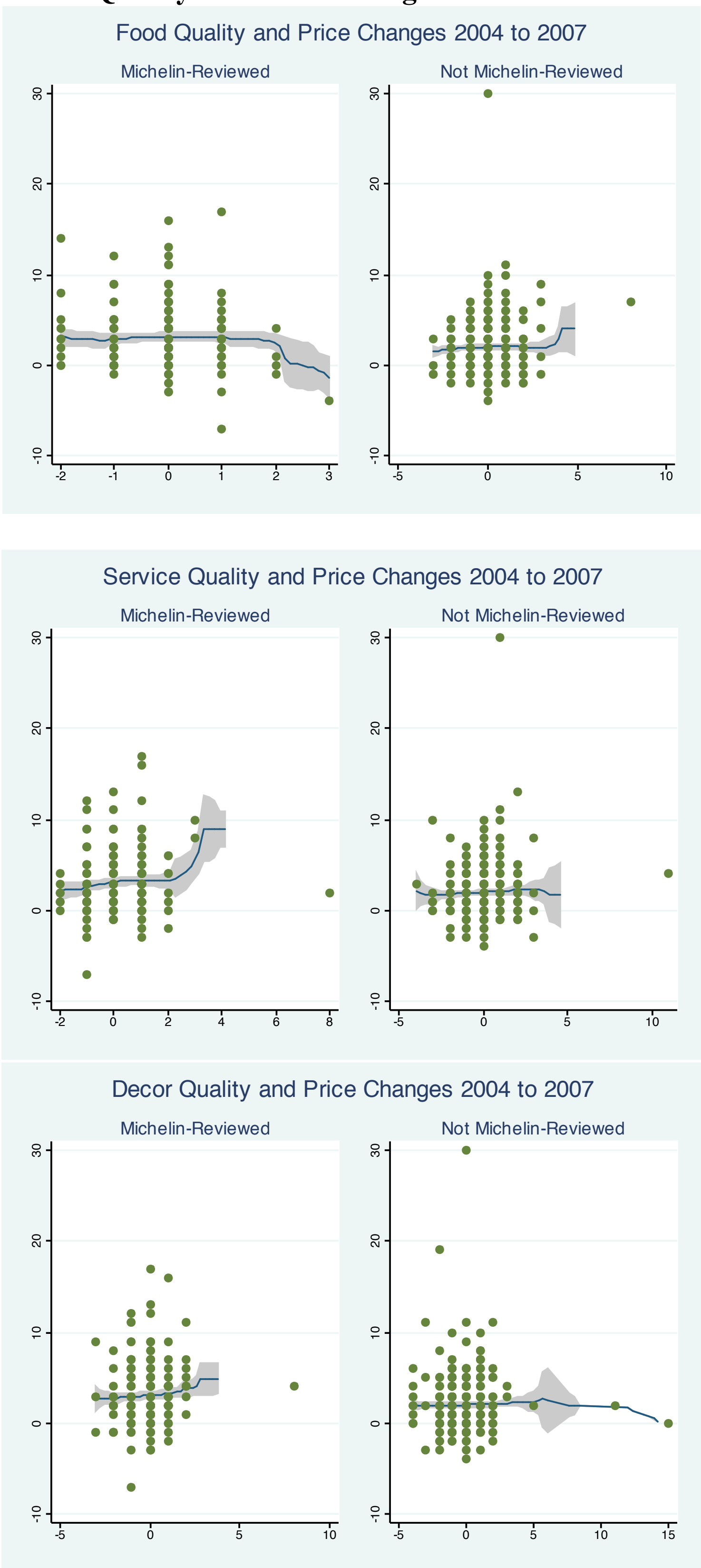


Figure 3

Quality and Price Changes 2004 to 2007



Price change in US$ (y-axis) and quality change in Zagat points (x-axis); Observations refer to restricted sample, i.e., when changes in the other quality variables are one or less (see text); the fitted line is the local polynomial smooth with the 95% confidence interval.