REPUTATIONAL INCENTIVES FOR RESTAURANT HYGIENE*

Ginger Jin
Department of Economics
University of Maryland
jin@econ.umd.edu

Phillip Leslie
Graduate School of Business
Stanford University
pleslie@stanford.edu

October, 2004

*We thank Dan Ackerberg, Lanier Benkard, Mara Lederman, Jon Levin and Joel Waldfogel. We are especially grateful to Steve Tadelis for his advice. We also thank the Los Angeles County Department of Health Services for allowing us to access the data. Financial support for this research was generously provided through NSF grants SES-0112295 (Jin) and SES-0112344 (Leslie).

1 Introduction

A theoretical literature emphasizes the potential of reputations for overcoming the problems associated with informational asymmetries. In a reputation mechanism, consumers may not observe product quality before making a purchase, but they learn from experience and form beliefs about product quality. These beliefs define firms' reputations. Reputations may have have two related but distinct effects. First, reputations can help consumers to make informed choices when there is a lack of information. Second, reputations may cause firms to provide higher unobserved product quality. Prior empirical studies of reputational incentives focus on identifying the first of these effects. Since consumer learning, or the demand-side effect, is the source of the incentives for firms, these studies are suggestive of the existence of the supply-side effect. In this paper we take a more direct approach to identifying the impact of reputational incentives on firm behavior. Using data on restaurant hygiene inspections, we perform a range of tests for whether reputational incentives are an effective determinant of restaurant hygiene quality.

Our data cover restaurant hygiene inspections in Los Angeles from July 1995 to December 1998. The inspections are conducted by Los Angeles County Department of Health Services (DHS) officials. Starting from the maximum score of 100, inspectors deduct specified points for each hygiene violation. Central to our approach is the fact that inspection scores before 1998 were not available to consumers, as they were for internal DHS use only. Consumers could request to see the list of violations at individual restaurants, but anecdotally we know this was rarely done. Following a hidden-camera television news expose of unsanitary restaurants, since the beginning of 1998 restaurants in Los Angeles are issued with hygiene grade cards—a letter grade (A, B or C) to be prominently displayed in the window, based on the score from their last inspection.

In Jin and Leslie (2003) we show that the introduction of restaurant hygiene grade cards in Los Angeles county in 1998, caused restaurants to increase hygiene quality, reducing foodborne illness hospitalizations by 20%. The impact of the grade cards indicates that consumers were lacking in information about restaurant hygiene before 1998. However, about 25% of inspections resulted in the equivalent of A-grade hygiene even before the grade cards.[1] If consumers are relatively uninformed about hygiene quality, why do these restaurants maintain such good hygiene? Our null hypothesis is that some restaurants are able to develop a reputation for good quality hygiene. Under this explanation, reputation formation varies across restaurants due to chain af-

[1] We show that restaurant fixed effects explain more than 60% of the variation in pre-grade card inspection scores, indicating that good quality hygiene is not due to transitory shocks, but rather systematic differences across restaurants.

filiation and/or regional differences in the degree of consumer learning, such as the prevalence of repeat-customers. We consider two alternative explanations: regional differences in willingness-to-pay for hygiene quality, and exogenous heterogeneity in restaurant hygiene quality (e.g. some managers like having a clean restaurant).

As mentioned above, prior empirical studies of reputation mechanisms focus on identifying the source of reputational incentives—the impact of consumer learning about firms' unobserved quality on demand.[2] Borenstein and Zimmerman (1988) study the incentives for airlines to maintain unobservable product safety. They find a small negative effect of a plane crash on the airline's demand, suggesting weak reputational incentives. However, as the authors note, the relationship between demand responsiveness to new information and the strength of reputational incentives depends on consumers' prior beliefs. A mild response by consumers to news of a plane crash may imply that either (i) consumers are unresponsive and therefore provide weak reputational incentives, or (ii) consumers have a strong prior belief that includes the expectation there will be an occasional crash, making them unresponsive to a single crash, but still providing a source of strong reputational incentives. Hence, the challenge in identifying the demand-side component of reputational incentives, is to analyze a dramatic piece of new information, such a several plane crashes within a short period of time, or to adequately control for consumers' prior beliefs. Hubbard (2002) examines whether firms that provide vehicle emissions tests are able to develop reputations for exerting high effort in helping vehicles pass. In the demand specification, Hubbard distinguishes the impact of an individuals past experience at a particular firm from the firm's failure rate, the latter serving as a proxy for consumers' prior beliefs. In the absence of a model of consumer learning, it is unclear if this approach adequately controls for consumers' prior beliefs.[3]

By analyzing the supply-side implications of reputational incentives, we avoid the complications associated with disentangling consumers' prior beliefs. Our approach is also conceptually straightforward. We test whether characteristics that give rise to a high degree of consumer learning, cause firms to exert higher unobservable effort. Two key features of the dataset facilitate our analysis. First is that we, the econometricians, observe the outcomes of restaurants' hygiene inspections, which are unobserved by consumers.[4] This allows us to analyze the relationship between hygiene quality and factors that impact the degree of consumer learning. However, the observed factors that affect consumer learning may be correlated with unobserved

[2] A number of papers study the role of reputations in eBay auctions. Most of these papers take the approach of estimating the effect of eBay seller ratings on the probability of sale and the final winning bid. Resnick and Zeckhauser (2002) and Bajari and Hortacsu (2003) provide detailed summaries of this literature. An exception is the study by Cabral and Hortacsu (2004), which analyzes the effect of eBay reputation on seller behavior.

[3] Chernew, Gowrisankaran and Scanlon (2002) estimate a fully-specified model of consumer learning, allowing them to measure the impact of new information on consumer behavior.

[4] It is assumed throughout this study that the DHS hygiene inspection score is an objective and meaningful measure of restaurant hygiene quality.

restaurant characteristics that also impact hygiene quality, such as the marginal cost of hygiene. To address this concern, we utilize a second key feature of our dataset: the introduction of restaurant hygiene grade cards. This policy change provides exogenous variation in the provision of information to consumers, allowing us to isolate the effect of consumer learning on hygiene quality, from other unobserved factors.

We also observe whether each restaurant is a company-owned or franchised chain unit. If chain affiliation is a source of reputational incentives, then we should observe higher average pre-grade card hygiene scores at chains than non-chains, all else equal. A finding that franchised units tend to have lower scores than company-owned units would indicate the presence of franchisee free-riding on chain reputation, providing further evidence that chain affiliation is a source of reputational incentives.[5] For independent restaurants we focus on local region factors as a source of reputational incentives for hygiene quality, such as the degree of repeat-customers. This implies regional clustering in hygiene scores. The grade card policy change allows us to distinguish local consumer learning effects from other local factors that also impact hygiene quality.

In Section 2 we describe the reputational incentives hypothesis and alternative explanations. Section 3 provides a review of the DHS inspections, grade card policy change, and summary statistics for our dataset. Basic evidence concerning the presence reputational incentives is presented in Section 4. Section 5 contains our analysis of whether chain affiliation is a source of reputational incentives. In Section 6 we test whether localized consumer learning is a source of effective reputational incentives. The conclusion is in Section 7.

2 Alternative Hypotheses

In the absence of restaurant hygiene grade cards, why do some restaurants have good quality hygiene? What explains why some restaurants consistently have low hygiene scores, while other restaurants have consistently high hygiene scores? In this section we describe our null hypothesis (reputational incentives) and a couple of alternative hypotheses.

Reputational Incentives Hypothesis

The reputational incentives hypothesis is that some restaurants have good quality hygiene because their customers can learn about their hygiene quality, allowing these restaurants to form

[5]Prior empirical studies of franchising include Brickley and Dark (1987), Norton (1988), Lafontaine (1992), and Lafontaine and Shaw (1999).

reputations for good hygiene, and obtain higher profits from doing so. Meanwhile, restaurants with low quality hygiene face customers that are unable to learn about their hygiene quality, preventing these restaurants from obtaining any benefit from providing good quality hygiene.

Two factors affect the formation of individual restaurant's hygiene reputation. First, local customers can learn about a restaurant's hygiene quality by repeatedly patronizing the restaurant, by talking to friends who have patronized the restaurant, or through exposure to local news reports about the restaurant. Brickley and Dark (1987) propose a prime example of when we expect there to be a low degree of consumer learning: for restaurants located near freeway exits, there are relatively few repeat-customers, leading to weak reputational incentives.[6] Whether the key feature is the degree of repeat-customers, or some other factor affecting consumers' ability to update their beliefs about hygiene quality, these factors are region-specific. All else equal, two restaurants located beside each other face similar consumer learning. This implies geographic clustering in the magnitude of restaurants' information differences.

Second, chain affiliated restaurants share the reputation of the chain as a whole. Customers of individual chain restaurants learn about hygiene quality for all restaurants in the chain, even if there are few repeat customers at each single unit. For example, a customer who has a bad experience with one restaurant in a chain may infer similar hygiene quality for all restaurants in that chain. We therefore expect that chain restaurants tend to face stronger reputational incentives than independent restaurants.

There is, however, an important distinction between company-owned chain restaurants and franchised chain restaurants. The owners (or investors) of a restaurant chain seek to maximize the sum of the profits of all company-owned units. The owner of a franchised chain restaurant seeks to maximize the profit of that unit alone. Combine this difference with the fact that consumers do not know if individual chain restaurants are company-owned or franchised, and we expect franchised chain restaurants to free-ride on a chain's reputation for good hygiene.[7] Hence, reputational incentives are lower for franchised chain restaurants than for company-owned chain restaurants, all else equal.

While franchised chains face lower reputational incentives than company-owned chains, there are a couple of reasons why franchised units may actually have higher quality hygiene than company-owned units. One reason is that it may be less costly for the owner of a franchised restaurant to monitor the hygiene effort of the employees than it is for the owner of a chain. This is because the owner of a franchised restaurant is often also the on-site manager, while

[6] Stigler (1961) also refers to tourists as a class of consumers lacking in knowledge about local markets.
[7] See Rubin (1978).

investors in a chain may be a diffuse group of stockholders.[8] A second reason is that the chain management chooses whether a particular unit is franchised, and would presumably take into account the degree of consumer learning in the local region of the restaurant when making this decision.

To summarize: under the reputational incentives hypothesis, variation in hygiene quality across restaurants is at least partly explained by informational differences across restaurants, which may be either region-specific and/or restaurant-specific. The former depends on the learning environment of the local region. The latter depends on whether a restaurant belongs to a chain and whether a chain unit is franchised or company-owned.

Alternative Explanations

There may be other regional factors that impact restaurants' hygiene quality, unrelated to informational differences. In particular, there could be regional differences in consumers' willingness-to-pay for hygiene quality.[9] There may also be restaurant-specific heterogeneity, also unrelated to informational differences, that affects hygiene. For example, restaurants may be heterogeneous in their hygiene costs. This explanation also requires consumers to possess some degree of information about restaurants' hygiene, or even low hygiene cost restaurants will have no benefit to providing good hygiene. But there need not be informational differences across restaurants. In this explanation, the cost of hygiene effort is an exogenous characteristic of each restaurant. A similar explanation could be that restaurant managers/owners are heterogeneous in their hygiene preferences—some managers like having a clean restaurant. As with the hygiene cost heterogeneity, this explanation is based on exogenous restaurant-specific heterogeneity. For our purposes, we merge these two explanations together, along with any other hypothesis based on exogenous restaurant-level heterogeneity, and we label this class of explanations as *exogenous restaurant heterogeneity*.

If there is exogenous restaurant hygiene, and consumers have imperfect information, then there may also be learning by consumers about restaurants' hygiene quality. While consumers' choices would be impacted by their learning, restaurants' hygiene quality is unaffected, as it is assumed to be exogenous. In other words, restaurants may obtain reputations for hygiene quality, but this does not cause any restaurant to improve their hygiene quality. Based on this reasoning, it can be said that learning is a necessary by not sufficient condition for effective reputational incentives.

[8] Caves and Murphy (1976) put forward this argument.

[9] There may also be regional differences in the cost of providing good hygiene, although this seems less plausible to us. Nevertheless, it would make no difference to include this in our approach.

However, there may still be sorting of restaurants across regions. For example, restaurants (or managers) with exogenously good hygiene quality, may choose to locate in regions where there is a high degree of consumer learning. Thus, regions with a high degree of consumer learning may also tend to have restaurants with good quality hygiene, even though hygiene quality is an exogenous restaurant characteristic. We consider this to also be the result of effective reputational incentives, and we make no attempt to distinguish this type of sorting effect from the impact of learning on restaurant hygiene effort. For this reason, the alternative hypothesis of exogenous restaurant heterogeneity may be better labeled as *exogenous restaurant heterogeneity (without sorting)*.

To summarize, the alternative hypotheses for the variation in hygiene quality across restaurants are: (i) regional differences in willingness-to-pay for hygiene quality, and (ii) exogenous restaurant heterogeneity (without sorting).

3 Data Summary

The data cover every restaurant inspection conducted by Los Angeles County DHS inspectors between July 1995 and December 1998. In California each county determines their own approach to conducting restaurant health inspections. Some counties, like Los Angeles, implement a scoring system, as an explicit attempt to reduce the impact on inspection outcomes of inspectors' subjective evaluations. Los Angeles inspectors deduct pre-specified points for each violation that is detected. For example, a food temperature violation results in a 5 point deduction, evidence of cockroaches results in a 3 point deduction, and a functioning but unclean toilet results in a 2 point deduction. For our purposes, we assume the DHS inspection scores are an objective measure of restaurants' hygiene quality.[10]

There are two changes in the score criteria that occur during our sample. Prior to July 1, 1997, inspectors could deduct up to 25 additional points based on their overall subjective evaluation of the restaurant's hygiene. This component was abolished in July 1997, leaving only the pre-specified point deductions for each violation. We can exclude the subjective deductions, allowing us to check the robustness of our findings. The second assessment change is less significant. In March 1998 the food temperature violation was divided into major and minor food temperature violations. In both cases, we presume that the average effect of the change in criteria on observed scores is a nominal change in inspection scores, with no real change in hygiene quality. In our analysis, we control for these changes by including dummy variables to capture

[10]To the extent that inspectors subjectivity is still a factor in determining scores, this implies measurement error in our dependent variable.

differences in mean scores due to the different criteria.

An important policy change applies to the final year of our data, which we exploit for some hypothesis tests. Beginning January 16, 1998, at the end of every inspection, restaurants are issued with a grade card to be prominently displayed in the window, near the entrance, for customers to see.[11] An A-grade is given for scores above 90, a B-grade for scores in the 80s, a C-grade for scores in the 70s, and for scores below 70 the numerical score is shown on the card. We assume the introduction of restaurant hygiene grade cards is an exogenous policy change, and in Section 4 we explain how this variation is helpful for identifying the presence of reputational incentives. The reason we believe the exogeneity assumption is reasonable, is because the policy change was prompted by a hidden-camera expose of unsanitary restaurants, by a local television news channel.[12]

The focus of this study is analyzing the determinants of restaurants' hygiene in the absence of hygiene grade cards. What power do the DHS inspectors have to force restaurants to maintain good quality hygiene? In the absence of grade cards, inspectors have almost no power. There are no fines for hygiene violations. Inspectors may close restaurants, but this is only in extreme cases, such as a fire or infestation, or if a restaurant gets a score below 60 in two consecutive inspections. Even then, the restaurant is closed only for the period of time it takes to rectify the problem (usually only a matter of days). Hence, a restaurant could consistently violate numerous hygiene standards, resulting in scores barely above 60, without incurring any kind of penalty. Inspectors educate restaurants' staff about hygiene safety, and try to convince them to make improvements, but ultimately have almost no power to assure compliance.

All of the tests we propose for identifying the presence of effective reputational incentives rely on the assumption that, prior to the grade cards, the results of the DHS inspections were not observed by consumers. Each week the *Los Angeles Times* newspaper reports the names of restaurants closed by the DHS. But as we noted, closures reveal a fraction of all hygiene violations. Restaurants were always required to provide the latest inspection report to any consumer that requested it. While we have no formal evidence concerning the extent to which consumers made such requests, we are confident this was sufficiently rare. The findings in Jin and Leslie (2003) support the claim that there is a general lack of information prior to the grade

[11] In fact, as we detail in Jin and Leslie (2003), the posting of grade cards is mandatory in some cities within LA county, and voluntary in other cities for an initial period before becoming mandatory. In both cases, grade cards are issued. The only difference is whether the manager has discretion over posting. We show in Jin and Leslie (2003) the effects on hygiene quality are similar in each case. We therefore abstract from this feature of the policy change in this study.

[12] Furthermore, the response of the regulators to the news story was fairly immediate. The story was aired on November 17, 1997. The county board of supervisors voted to implement grade cards on December 17, 1997. Inspectors began issuing grade cards on January 16, 1998. See Jin and Leslie (2003) for more details of the policy change.

cards.

We observe the name and address of each restaurant. This allows us to associate local demographic data from the census with each restaurant, as well as information on local businesses (such as the number of hotel employees working in the same zip code). From restaurant names and the Yellow Pages we can identify cuisine type for approximately half of the restaurants. We also obtained the Zagat Survey restaurant guide for each of the corresponding years in our data. From Zagats, we identify which restaurants are included in the guide, and their associated review scores. Restaurant names also allow us to identify chain restaurants.[13] We can further distinguish company-owned chain units from franchised units, on the basis of ownership information, also provided by the DHS. Although the data does not include a variable that indicates if each chain restaurant is franchised, from the name of the owner we can infer the type of ownership. Basically, we distinguish owners that are company-names from owners that are names of individuals. For names that are ambiguous, such as "Licensing Department", we classify them as company-owned. In doing so, we are more likely to be biased towards underestimating the impact of franchise ownership.

Table 1 provides a summary of the score data, by different types of restaurants, and distinguishing pre- and post-grade card scores. There are 24,304 restaurants that were inspected a total of 127,111 times. The mean score for all pre-grade card inspections is 76.77, compared to the post-grade card mean of 89.62. The dispersion of the unconditional score distribution is much less after grade cards. Less than 5% of restaurants are covered by the Zagats guide. These are undoubtedly the more fancy (and expensive) restaurants in the data. As one might expect, the average pre-grade card hygiene score of Zagats restaurants is above the average for all restaurants. After grade cards, Zagats restaurants have slightly below average hygiene scores.

There are 2,632 chain restaurants, equal to nearly 11% of restaurants in the data. Chain restaurants have significantly higher average hygiene scores than non-chain restaurants, before the grade cards. After grade cards, chains continue to have better hygiene, although the difference is reduced by about half. About 63% of chain restaurants in our data are company-owned units. Before the grade cards, the average score for company-owned chain units is about one point higher than franchised units. The difference in average hygiene scores between company-owned and franchised chains, after grade cards, is negligible.

In Table 1 we also show average hygiene scores for each of the top 6 chains, distinguishing between company-owned and franchised units. There is noticeable heterogeneity across chains in the fraction of units that are franchised, ranging from 20% for El Pollo Loco to 57% for McDonalds. There is also variation in average hygiene scores across chains. Comparing average

[13] We use *Bond's Franchise Guide* to identify national and regional chains.

pre-grade card scores for the company-owned units, the highest is Burger King (86.98) and the lowest is McDonalds (81.09). Although it should be noted that the average for McDonalds is still 4.32 points higher than the average for all restaurants. While company-owned Burger King units have relatively high average hygiene quality, before grade cards, the franchised Burger King units have significantly lower average scores. The average score of franchised units of Burger King is 4.89 points below the company-owned units of the same chain. Burger King is not alone on this count. The average score of franchised units of El Pollo Loco is 4.91 points below the company-owned units. At the other end of the spectrum, franchised McDonalds units have 0.69 higher average scores than company-owned McDonalds, before grade cards.

Table 1 also shows that after grade cards, there are smaller differences in average hygiene between company-owned and franchised units of the same chain. For example, after grade cards, the difference in average scores between company-owned and franchised Burger King restaurants reduces to 0.1. KFC is an even more striking example. Before grade cards the franchised units are 3.37 points below the chain units, on average. After grade cards, the franchised units of KFC are 1.21 points above the company-owned units, on average.

We also report mean scores for several types of cuisine. Burger, chicken and pizza restaurants all tend to have above average hygiene. Chinese and Mexican restaurants have below average hygiene. Chinese restaurants have the lowest average hygiene scores before grade cards: 70.68, or 6.09 points below the average of all restaurants. However, Chinese restaurants show significant improvement after grade cards, increasing to a mean of 86.13, or 3.49 points below the average of all restaurants. The DHS data also records the number of seats in each restaurant (as a categorical variable), and Table 1 also shows conditional means according to the number of seats.

In the last two rows of Table 1 we compare hygiene on the basis of per-capita income in census tract of each restaurant. A restaurant is designated to be in a lower income area, if per-capita income in their census tract is below the median for Los Angeles county. Otherwise, the restaurant is designated to be in a higher income area. The results are quite provocative. Before grade cards, the average hygiene score of restaurants in lower income areas is 4.24 points below restaurants in the higher income areas. In contrast, after grade cards, the average score of restaurants in the lower income regions is slightly above the average for the higher income regions. On the face of it, this suggests the introduction of grade cards has an even bigger positive impact on restaurant hygiene in poorer neighborhoods.

In this section, we have described some of the salient features of restaurant inspections in Los Angeles, including the introduction of grade cards, which serves as an important source of variation for some of our tests. We have also presented basic summary statistics of the data, with

an emphasis on the differences between chains and non-chains, as well franchised chain units and company-owned chain units. This variation also plays an important role in some of our hypothesis tests. In the next section, we present some basic evidence in support of our general approach. Then in the following sections we explain and implement a variety of hypothesis tests, designed to identify the presence of effective reputational incentives.

4 Basic Evidence

In this section we present basic evidence in support of our approach to identifying an effect of reputational incentives on restaurants' behavior. An important premise of the reputational incentives explanation is that variation in hygiene scores is due to systematic differences across restaurants. These systematic differences may be related to restaurant's characteristics, or characteristics of their local neighborhood. An alternative is that the score variation is due to inspectors' idiosyncracies (despite the use of a score-based assessment criteria), or due to restaurants incurring hygiene shocks over time. Since we observe each restaurant inspected multiple times, we perform a variance decomposition to help gauge the relative importance of these different factors.

Table 2 shows the results of several variance decompositions. In the top panel, an observation is a restaurant inspection before the introduction of grade cards. This includes inspections under regime I (objective plus subjective components) and inspections under regime II (no subjective component). In the middle and bottom panels of Table 2 we report the results using only regime I, and depending on whether we include the formal subjective component.

As reported in Table 2, there are 83,790 inspections before the grade cards in our data. Conditioning the observed scores on quarterly dummies and inspection regime dummies only, explains 4% of the score variation. Our data include 38 observable restaurant characteristics (such as number of seats, whether included in Zagats, whether a chain and so forth), and also conditioning on these variables explains 11% of the score variation. However, if instead of conditioning on observed restaurant characteristics, we condition on all time-invariant restaurant characteristics (ie. restaurant fixed effects) then we explain 62% of the score variation. This is basic evidence in support of the claim that pre-grade card hygiene scores are largely due to systematic differences across restaurants.

Some of our tests also focus on identifying differences in the degree of consumer learning across regions. If regional differences in reputation formation are a determinant of restaurants' hygiene quality, then we expect region fixed effects to also explain a sizable fraction of score

variation. This would reflect the fact that two restaurants near each other are more likely to face similar incentives for hygiene quality, than two restaurants far apart. In the top panel of Table 2, we report that city-level fixed effects (there are 151 cities in Los Angeles county during this period) explain about 20% of the variation in scores. Five-digit zip code fixed effects (of which there are 315) explain 27% of the variation. This suggests that local region characteristics may explain up to 43% of the systematic differences in restaurants' hygiene qualities.[14] Note also that city fixed effects, with less than half the number of variables than zip effects, explain nearly as much of the score variation as the zip fixed effects. Of course region fixed effects include factors other than the degree of learning by local consumers. Nevertheless, this is basic evidence in support of the potentially important role that local factors may play in determining restaurant hygiene.

In the middle panel of Table 2 we report results when we limit the analysis to inspections only under the first regime (66,977 observations). In this case, the constant and quarterly dummies alone explain effectively zero of the score variation, indicating the absence of any seasonality in hygiene quality, whatsoever.[15] Restaurant fixed effects now explain 69% of the score variation, which is slightly more than in the top panel. The remaining 31% of unexplained score variation may be due to inspectors' subjectivity or shocks to restaurant hygiene. To the extent that inspectors' subjectivity is a significant factor, we expect this would be particularly important in the determination of the formal subjective component of the assessment. In the bottom panel of Table 2 we decompose the variance of the scores under regime I, with the formal subjective component subtracted out. Restaurant fixed effects now explain 71% of the variation, compared to 69% when we include the subjective component. Such a small difference suggests that inspectors' idiosyncracies are not a major factor in the determination of hygiene scores, although it is not conclusive.

A simple approach to examining the effect of reputational incentives on restaurants' hygiene quality would be to estimate an OLS regression on a cross-section of restaurants, in which the dependent variable is pre-grade card hygiene inspection scores. Regressors would include variables that capture the degree of consumer learning, and variables that control for other factors which may impact hygiene quality. One could then test whether the coefficients on the learning variables are significantly positive. However, there are two problems with this approach. First, it is difficult to obtain convincing measures of the degree of consumer learning about restaurants. Second, there is good reason to expect consumer learning about restaurants is correlated with other factors that impact on hygiene quality, some of which are invariably

[14] Restaurant fixed effects explain 62% and zip fixed effects explain 27%. Suggesting the possibility that zip characteristics account for up to 43% (27/62) of the cross-restaurant variation in scores. This is of course merely suggestive.

[15] In the top panel, the constant explains 4% of the variation, which is presumably all due to the regime dummy that is included in that specification.

unobserved by us. For example, in regions where consumers tend to have a high willingness-to-pay for hygiene quality, consumers may be more likely to obtain information and learn about hygiene quality at their local restaurants. Nevertheless, the results of such a regression are shown in Table 3. The regression is estimated using pre-grade card inspection only. We include as many restaurant characteristics as possible to mitigate potential bias due to missing variables. In the table we report coefficient estimates for the most interesting variables, leaving off several that are listed in the notes to the table.

Recall that in Table 1, the average hygiene score of company-owned chain restaurants is 6.2 points above the average for all restaurants. In the regression, which controls for numerous other factors, this difference is now estimated to be 4.3. Similarly, Table 1 shows the average score of franchised chains is 1.1 less than the average score of company-owned chains. The estimated difference reported in Table 3 is .6, and is significantly different from zero with 95 percent confidence. We infer from these findings that chain affiliation, and the tendency towards franchising, are correlated with other observed factors that also predict lower hygiene scores. The *Zagats restaurant* variable is a dummy for whether the restaurant is included in the Zagat Survey. This could be interpreted as a proxy for whether a restaurant has a strong reputation. However the guide does not include information on hygiene quality, so this is unlikely to be a direct measure of a restaurant's reputation for hygiene quality. Nevertheless, it is conceivable that reputation for food and service quality is correlated with reputation for hygiene quality. Consistent with this interpretation, in Table 3 we report the estimated coefficient is 1.8 (significantly different from zero with 95% confidence). But we also find the numerical food rating that appears in Zagats (*Zagats food rating*) is negatively related to hygiene scores, casting doubt on this interpretation.

The main variables of interest in Table 3 are the proxies for the degree of repeat-customers. As mentioned above, Brickley and Dark (1987) utilize a dummy variable for whether a restaurant is located near a freeway exit, as an indicator that the restaurant has few repeat-customers. While this is appealing for restaurants located outside a metropolitan region, we expect this would be a poor measure for restaurants in Los Angeles County. Instead, we propose two kinds of proxies for the degree of repeat customers. First are measures of tourist activity in the local area. Using the zip code business pattern data from the Census Bureau, we observe employment by industry and zip code, for each year 1995 to 1998. As measures of tourist activity, we specifically use the fraction of total employment in each of hotels, museums and amusement parks. The logic is: the greater the fraction of employees in each of these industries in a given zip code, the fewer repeat-customers for restaurants, due to the increased number of tourists. The estimated coefficients for these variables are shown Table 3. Given our logic, we expect these variables to have negative coefficients. However, the estimated coefficients are positive or insignificant.

The second set of measures for the degree of repeat-customers are based on revealed-preference arguments. Assuming that chain restaurants have an advantage over independent restaurants because of their chain reputations, we expect that chain restaurants are more likely to open in locations with relatively few repeat customers. We define the variable *mostly chain restaurants in zip* as a dummy equal to one for restaurants located in zips where at least 15% of restaurants are chains (which is a quarter of the zips). By revealed preference, we expect these zips have relatively few repeat-customers. We therefore expect this variable to be negatively correlated with hygiene scores. Contrary to our expectation, in Table 3 we report the estimate to be a highly significant 1.8.

We also define the variable *mostly independent restaurants in zip* as a dummy equal to one if the percent of restaurants that are chains in the zip is less than 5% (which is also a quarter of the zips). In this case, we expect these are regions with a relatively high degree of repeat business, leading to higher average scores. But again in contrast to our expectation, the coefficient estimate is negative and highly significant.

Another revealed-preference measure of repeat-customers is the fraction of chain units that are franchised. Brickley and Dark (1987) conjecture that chain restaurants located near freeways are more likely to be company-owned rather than franchised. This is because of the relatively low degree of repeat-customers traveling along freeways, leading to a higher propensity of free-riding by franchised units in these locations. Brickley and Dark (1987) actually find the opposite is true—chain units near freeways are more likely to be franchised. Nevertheless, we also apply the logic that underlies the conjecture of Brickley and Dark (1987). Assuming that chain units are more likely to be company-owned in areas with many relatively few repeat-customers, we infer from the presence of a high ratio of company-owned to franchised units, that there is a low degree of repeat customers.

Following this logic, we define the variable *mostly company-owned chains in zip* as a dummy equal to one for restaurants located in zip codes where the fraction of chain restaurants that are company-owned is greater than 25% (which is over half the zips). We expect this variable to be negatively related to hygiene scores. Again we find the opposite. We also define the variable *mostly franchised chains in zip* to equal one for restaurants located in zips where fraction of chain restaurants that are company-owned is less than 50%. We expect the associated coefficient to be positive, because we interpret this as an indicator of a high degree of repeat-customers. Unsurprisingly by this point, the estimate is the reverse. As in Brickley and Dark (1987), we find the empirical relationships are at odds with the stated logic.

In the next section we show that chain affiliation seems to be a source of effective reputational incentives, causing chains to have significantly better hygiene than non-chains. We also find

evidence in support of the presence of franchisee free-riding, but the effects appear to be quite small in magnitude. This suggests that whether a chain unit is franchised or company-owned is of second-order importance to the effect of chain affiliation on reputational incentives. It is therefore conceivable that the first of the above revealed-preference measures—capturing the fraction of restaurants that are chain-affiliated in each region—is an a-priori better measure of reputational incentives facing the independent restaurants in the region, than the mostly-franchised proxy.

In Table 3 we also report the estimates for several demographic variables. The estimate on *Fraction married in census tract* implies a one percent increase in the number of married households is associated with a .18 increase in average restaurant hygiene. Similar interpretations attach to the other "Fraction" variables.

Instead of relying on the above regression findings to assess the presence of reputational incentives, we exploit a source of exogenous variation—the introduction of restaurant hygiene grade cards. Three aspects of this policy change are relevant for our purposes. First, grade cards provide exogenous intertemporal variation in the amount of information available to consumers about restaurants' hygiene quality. Second, grade cards eliminate any information differences across restaurants—the degree of repeat-customers or chain affiliation are less informative about hygiene quality than grade cards. Third, the grade cards have no impact on the cost of hygiene quality or consumers inherent willingness-to-pay for hygiene quality. Grade cards may cause restaurants to incur greater hygiene costs, but this is because they choose to incur these costs in the face of increased incentives. Grade cards have no impact on the hygiene quality cost function.

In this section we have provided basic evidence in support of the argument that hygiene quality (in the absence of grade cards) is largely a function of time-invariant differences across restaurants. Furthermore, the finding that city or zip code fixed effects can explain a significant fraction of the variation in scores, indicates that regional heterogeneity may be an important source of the time-invariant differences across restaurants. These facts are consistent with our hypothesis that reputational incentives are a determinant of restaurant hygiene quality. We also explored the correlation between hygiene scores and several possible proxies for the degree of repeat-customers in the local area of restaurants. However, we see no compelling reason to view this approach as identifying causal effects. In the next section we present formal hypothesis tests, which are designed to identify a causal effect of reputational incentives on restaurant hygiene quality.

5 Chain Affiliation as a Source of Reputational Incentives

As noted above, chain affiliation may be a source of reputational incentives for good quality hygiene. We therefore expect chain restaurants to have higher average hygiene quality than non-chains, before the introduction of grade cards. We also noted that franchised chain units may free-ride on the reputation of the chain, exerting less hygiene effort. Hence, a finding that franchised units have lower average hygiene scores than company-owned units would also be evidence in favor of the reputational incentives provided by chain affiliation. The challenge to testing both of these implications is being able to control for other factors that are correlated with chain affiliation and/or franchise-ownership, which also impact hygiene quality.

The regression results reported in Table 3 indicate that company-owned chains have better hygiene than franchised chains, which in turn have better average hygiene than non-chains. The regression includes controls for observed restaurant and region characteristics. But there may be unobserved region characteristics, such as the degree of consumer learning (or extent of repeat-customers) that are correlated with the chain and franchising variables. To address this concern, we include region fixed effects. Let s^B_{ijt} denote the hygiene inspection score obtained by restaurant i, in region j, in period t, before the introduction of grade cards (superscript B denotes "before grade cards"). Using only pre-grade card hygiene inspections, we estimate the specification

$$s^B_{ijt} = \alpha^B_j + \beta^B c_i + \gamma^B f_i + X_i\theta^B + \epsilon_{ijt}, \tag{1}$$

where α^B_j is a region-fixed effect, c_i is a dummy equal to one if restaurant i is a chain restaurant, f_i is a dummy equal to one if the restaurant is a franchised chain unit, and X_i are observed restaurant characteristics (including cuisine).

If chain affiliation is a source of reputational incentives, we expect $\beta^B > 0$ and $\gamma^B < 0$. These coefficients are identified by within-region variation in c_i and f_i, respectively. The results are reported in the first column of estimates in Table 4. The estimate for the chain coefficient is 4.9 and is significantly different from zero with 99 percent confidence. The estimate for the franchising coefficient is -0.6 and is significantly different from zero with 95 percent confidence. These estimates support the hypothesis that chain affiliation is a source of reputational incentives. We also estimate a version of equation (1) that includes separate chain dummies for each of the top ten chains. This ensures we identify the franchising coefficient from within-chain, within-region, variation in franchising (rather than cross-chain variation in franchising).

Since identification of the parameters of interest does not rely on an explicit source of exogenous variation in chain affiliation of franchising, there may be other differences between company-owned and franchised units that also impact on hygiene quality. The inclusion of re-

gion fixed effects precludes some kinds of potential biases, but perhaps not all. There could be variation in willingness-to-pay for hygiene quality or consumer learning across locations within a region, which may introduce bias. An even more serious concern is that chains may tend to have better hygiene because the cost of hygiene effort is lower at chains, unrelated to any effects due to reputational incentives. Including X_i helps to mitigate this. For example, chains are often hamburger restaurants, and hamburger restaurants may have lower hygiene costs than other kinds of restaurants. But clearly this is not an ideal solution.

To address these concerns, we exploit the grade card policy change. By increasing the provision of information to consumers about restaurants' hygiene, grade cards are a substitute for good hygiene reputations.[16] Hence, if the difference in average hygiene quality between chains and non-chains is reduced by the grade cards, me may infer the effectiveness of chain reputation to maintain good hygiene, in the absence of grade cards. One approach to testing this implication is to estimate the following specification, using the pooled data for before and after the grade cards:

$$s_{it} = \alpha_i + \beta_1 g_t + \beta_2 g_t c_i + \gamma g_t f_i + \epsilon_{it}, \tag{2}$$

where s_{it} is the inspection score at restaurant i in period t (including observations before and after the grade are introduced), g_t is a dummy equal to one for inspections occurring after the introduction of hygiene grade cards.[17] A virtue of this approach is the inclusion of restaurant fixed effects (α_i) to control for all time-invariant restaurant heterogeneity. Assuming the grade cards have no impact on the hygiene cost function or consumers' willingness-to-pay for hygiene, the approach allows us to isolate the informational effect of chain affiliation.

A finding that $\beta_2 < 0$ indicates the increase in average hygiene quality, due to the grade cards, is larger for non-chains than for chains. And a finding that $\gamma > 0$ implies the grade cards have a bigger positive impact on franchised chains than company-owned chains, indicating the presence of franchisee free-riding in the absence of grade cards. The results are reported in the second column of estimates in Table 4. The estimate for β_2 is -4.1, and the estimate for γ is 1.2. Both estimates are significantly different from zero with 99 percent confidence. These findings again support the hypothesis that chain affiliation is an effective source of reputational incentives for good quality hygiene.

Note also that we obtain an estimate for the stand-alone franchise variable in the second specification of Table 4, even though this model also includes restaurant fixed effects. Identification follows from instances of individual chain restaurants changing from being company-owned to franchised (or the reverse). Based on this variation in the data, we estimate a fairly large and

[16] Grade cards may be an imperfect substitute for hygiene reputations. The main point is that grade cards reduce the informational differences for consumers between chains and non-chains.

[17] The j subscript (indexing regions) is dropped, because the specification includes restaurant fixed effects.

significant negative effect of franchising on hygiene (-2.3). We conclude that there is evidence of some degree of franchisee free-riding, but the evidence is mixed on whether the effect on hygiene quality is large or small in magnitude. The evidence presented in Table 1 suggests that franchisee free-riding may be a major problem for certain chains, and this is probably why the estimated effects differ depending on which source of variation is relied on for the estimates.

A concern with the interpretation of the negative estimate of β_2 in equation (2), is that such a finding may be a due to a particular form of cost heterogeneity, rather than differences in pre-grade card reputational incentives. Suppose that (i) the marginal benefit of hygiene quality is a decreasing function of hygiene quality and is the same for chains and non-chains; (ii) grade cards cause an increase in the marginal benefit of hygiene; (iii) the marginal cost of hygiene is an increasing function of hygiene quality; and (iv) the marginal cost function of hygiene for chains is less than for non-chains, but the difference between the two is decreasing in hygiene quality. In this setting, it is possible that grade cards may cause the average chain to improve hygiene from 83 to 93, while the grade cards cause the average non-chain to improve hygiene from 77 to 89 (based on the mean scores in Table 1). This implies an estimate of $\beta_2 < 0$. But the result is due to cost heterogeneity, rather than any pre-grade card differences in reputational incentives.

To address this concern, we estimate the following specification, as a robustness check of our earlier findings:

$$s_{ijt}^{B} = \alpha_j + \beta c_i + \gamma f_i + \delta \bar{s}_i^{A} + X_i\theta + \epsilon_{ijt}, \tag{3}$$

where $\bar{s}_i^{A}$ is the average post-grade card inspection score for restaurant i.[18] Conditional on obtaining the same hygiene score after grade cards, β captures any systematic difference in pre-grade card scores between company-owned chains and independent restaurants. If the relative impact of the grade cards is greater for restaurants with lower initial scores, and is unrelated to chain affiliation, then we should obtain estimates of $\beta = 0$. Hence, a finding that $\beta > 0$ would indicate that the shortage of information of before the grade cards, causes non-chain restaurants to have lower quality hygiene than chain restaurants. Similarly, the coefficient γ capture the difference in pre-grade cards scores between franchised and company-owned chain restaurants, while controlling for post-grade card scores. If there is free-riding, we should find $\gamma < 0$.

The interpretation of the estimates in equation (3) is further strengthened with the assumption that the grade cards equalize informational differences across restaurants. Given this assumption, variation in post-grade card scores across restaurants, must be due to differences in: (i) the marginal cost of hygiene quality, and/or (ii) consumers willingness-to-pay for hygiene quality. By including region fixed effects in equation (3), it is assumed that we control for differences in consumers' willingness-to-pay for hygiene quality. Hence, by assumption, within-region

[18] To compute the average we control for the minor inspection criteria change in March 1998.

variation in post-grade card scores is because of heterogeneity in the marginal cost of hygiene quality.[19] If two restaurants in the same region have the same post-grade card hygiene score, then we infer they have the same marginal cost of hygiene function.[20] If these two restaurants have different hygiene scores before grade cards, it must be due to informational differences which affect the marginal benefit of hygiene quality. An estimate of $\beta > 0$, therefore, implies that chain affiliation is a source of informational differences, creating an incentive to maintain good quality hygiene.

In Table 4 we report two sets of estimates for the coefficients in equation (3), that vary according to whether the region fixed effects are included (α_j). The only difference is whether the estimate for γ is significantly different from zero. In the third column of Table 4 we report the estimates for the specification without region fixed effects. In this case, the coefficient on the chain variable is estimated to be 5.4, and the estimate on the franchising variable is -1.7. Both are significantly different from zero with 99 percent confidence. When we include city fixed effects, the chain effect decreases to 3.8 and the franchising effect is insignificant, as shown in the fourth column of estimates.[21]

In the presence of city fixed effects, the franchising coefficient is identified by within-city variation in franchising, while controlling for post-grade card scores. The ideal variation is to have two chain restaurants in the same city with the same post-grade card scores, but one is franchised and the other company-owned. Perhaps not surprisingly, there is relatively little of this kind of variation in the data, leading to insignificance in the estimate of γ. Nevertheless, we obtain a high degree of significance on the chain coefficient.

The estimate for δ in equation (3) is about .5 and is highly statistically significant. This implies that lower post-grade card scores, are associated with smaller improvements in hygiene from their pre-grade card level.[22] The concern that motivated the specification of equation (3) was that restaurants with low quality before grade cards, would find it easier to improve their hygiene after the grade cards, than restaurants which already have good hygiene quality. This would imply an estimate of $\delta > 1$. Instead, the finding that $\delta = .5$ indicates the reverse—it is harder for restaurants with low pre-grade card hygiene quality to improve, than it is for

[19] If we assume that marginal cost and marginal benefit are linear in hygiene scores, but not necessarily equal for chains and non-chains, then it can be shown that $\delta \bar{s}_i^A$ fully controls for cost heterogeneity.

[20] We also assume that marginal cost of hygiene curves for different restaurants never cross. This rules out the possibility that the marginal cost of hygiene at one restaurant may be lower than another restaurant for low hygiene levels, but is greater at high hygiene levels. Which seems reasonable. Note, we do not assume that marginal cost functions are parallel shifts across restaurants.

[21] The results are unchanged if we define regions as zip codes. If we narrow the region definition to the census tract level, the qualitative aspects of the results are robust, but some estimates become statistically insignificant.

[22] For simplicity, rewrite equation (3) as $s^B = .5s^A$. This implies $(s^A - s^B) = .5s^A$. Hence, the change in hygiene is increasing in s^A.

restaurants with good hygiene.[23] More specifically, it appears that the difference in the marginal cost of hygiene between restaurants with the same score, tends to increase at higher levels of hygiene. In any event, the estimates for this specification reinforce our conclusion that chain affiliation is an effective source of reputational incentives for good quality hygiene.

In this section we have tested the hypothesis that chain affiliation is source of reputational incentives that causes chain restaurants to maintain good quality hygiene. The summary statistics in Table 1 suggested this may be true, but there could have been other explanations for the basic difference in mean scores between chains and non-chains. The analysis in this section indicates that the difference is indeed due to the effect of chain reputation. Our findings also reveal the presence of free-riding on chain reputation by franchised chain units. However, this effect seems fairly small in magnitude.

6 Local Region Reputational Incentives

The analysis of the previous section concerns reputational incentives related to chain affiliation. In this section, we examine evidence concerning reputational incentives for independent restaurants. If we observed a measure for the degree of repeat customers at each restaurant, it would be possible to analyze its effect on hygiene quality in a similar manner to the analysis of chain effects in the prior section, also drawing on the grade cards to disentangle spurious correlations. But we have no such measure for independent restaurants.

Instead, we rely on the assumption that the degree of consumer learning, through which reputation formation occurs, is a characteristic of the local region, if it exists at all. We noted in Section 4 that city fixed effects explain around 20% of the variance in pre-grade card scores. This suggests region factors play a significant role in determining restaurant hygiene quality, but not necessarily because of consumer learning in particular. There may also be regional variation in consumers' willingness-to-pay for hygiene quality. The following tests exploit the introduction of grade cards to disentangle these effects. We divide the analysis into three subsections. The first two parts are based on comparisons of regional fixed effects under alternative assumptions. In the third subsection we analyze the importance of regional factors in the reputation effects of chain affiliation.

[23] As a robustness check, we estimate a version of equation (3) that allows $\bar{s}_i^A$ to enter as a set of dummy variables. The result did not change.

6.1 Absolute Differences in Regional Effects

We separately estimate the following two equations, using inspections conducted before and after grade cards, respectively:

$$s^B_{ijt} = \alpha^B_j + \beta^B c_i + \gamma^B f_i + X_i\theta^B + \epsilon_{ijt}, \quad \text{and} \tag{4}$$

$$s^A_{ijt} = \alpha^A_j + \beta^A c_i + \gamma^A f_i + X_i\theta^A + \epsilon_{ijt}. \tag{5}$$

In this section we focus on the region fixed effects: α^B_j and α^A_j. The estimates of the region effects compound reputational incentives and willingness-to-pay for hygiene. Let r_j denote the degree of reputational incentives, or consumer learning, facing restaurants in region j. Let w_j denote consumers' willingness-to-pay for hygiene quality in region j. We assume the following functional form for the impact of r_j and w_j on average hygiene quality in region j:

$$\alpha^B_j = \alpha_1 r_j + \alpha_2 w_j + \alpha_3 r_j w_j, \quad \text{and} \tag{6}$$

$$\alpha^A_j = \alpha_1 g + \alpha_2 w_j + \alpha_3 g w_j. \tag{7}$$

In equation (7), g replaces r_j as both are measures of consumers' information, and we maintain the assumption that grade cards are at least as informative as any restaurants' reputation for hygiene quality.

Importantly, the formulation shown in equations (6) and (7) allows for: (i) separate effects from r_j and w_j, and (ii) an effect due to the interaction r_j and w_j, which may have either a positive or negative coefficient. Of course r_j and w_j are unobserved to us, the econometricians, preventing us from estimating α_1, α_2 and α_3. Nevertheless, our goal is to determine if r_j is a significant determinant of restaurants' hygiene quality in the absence of grade cards.

In this subsection, we assume that $\alpha_3 = 0$, ruling out any interaction effect between information and willingness-to-pay for hygiene quality. We relax this assumption later. The assumption allows us to test for the presence of reputational incentives by simply computing the difference between the before and after region fixed effects. Combining equations (6) and (7), and assuming $\alpha_3 = 0$, implies:

$$\alpha^A_j - \alpha^B_j = \alpha_1(g - r_j).$$

Hence, if $(\alpha^A_j - \alpha^B_j)$ is statistically different across regions, then r_j must also be statistically different across regions. Moreover, high values of $(\alpha^A_j - \alpha^B_j)$ imply relatively low values of r_j (assuming $\alpha_1 \geq 0$). Although we are unable to say anything about the absolute levels or r_j, since the level is confounded with g which is unobserved. An F-test supports the hypothesis that $(\alpha^A_j - \alpha^B_j)$ is not constant ($F = 38.00$) with 99 percent confidence.

All of our tests for local region reputation effects rely on the assumption that variation in the region fixed effects, after grade cards, is due to differences in willingness-to-pay, since the grade cards equalize consumers' information across regions. This particular test relies on a stronger version of this assumption—absolute differences in the region fixed effects after grade cards, measure the absolute differences in consumers' average willingness-to-pay for hygiene quality across the regions. Hence, by subtracting the post-grade card region fixed effects, α_j^A, from the pre-grade card region fixed effects, α_j^B, we can isolate regional differences in pre-grade card reputational incentives.

How reasonable is it to assume that $\alpha_3 = 0$? It is hard to say. On the one hand it could be argued that, even if $\alpha_3 \neq 0$, it is likely to be small in magnitude compared to the first-order effects of r_j and w_j on hygiene quality. On the other hand, it seems plausible to us that reputational incentives are more powerful for fancy restaurants, which would also be restaurants located in regions with a high willingness-to-pay for hygiene quality. Regardless, in the next subsection we propose a test that allows for non-zero values of α_3.

6.2 Relative Differences in Regional Effects

We now allow for the possibility that hygiene quality also depends on the interaction between reputational incentives and willingness-to-pay for hygiene quality ($\alpha_3 \neq 0$). Rearranging equation (6) yields

$$w_j = \frac{\alpha_j^B - \alpha_1 r_j}{\alpha_2 + \alpha_3 r_j}.$$

Substitute into equation (7), to obtain

$$\alpha_j^A = \left(\alpha_1 g - \frac{\alpha_1\alpha_2 r_j + \alpha_1\alpha_3 g r_j}{\alpha_2 + \alpha_3 r_j}\right) + \left(\frac{\alpha_2 + \alpha_3 g}{\alpha_2 + \alpha_3 r_j}\right)\alpha_j^B.$$

If $r_j = r$, then the above equation reduces to

$$\alpha_j^A = \kappa_1 + \kappa_2 \alpha_j^B,$$

where κ_1 and κ_2 are constants. Hence, if there is no regional variation in reputational incentives, then the post-grade card region fixed effects are an affine transformation of the pre-grade card fixed effects. We test if this is true.

Figure 1 helps provide a more intuitive explanation for the approach. The figure depicts the region fixed effects during three distinct periods of time. The first period is July 1995 to June 1996, and is shown along the horizontal axis. The second period covers July 1996 to June 1997 (the dots in the figure). And period three covers 1998, following the introduction of

grade cards (the crosses in the figure). The inspection regime is identical during the first two periods. It is therefore not surprising that the dots in Figure 1 are close to the 45 degree line. This serves as a robustness check. The crosses in Figure 1 show the relationship between the fixed effects in the first and third periods. Clearly the ordering of the fixed effects has been dramatically changed in the third period. This implies the relative effect of the grade cards is quite different across regions. We presume the biggest relative improvements occur in cities with a low degree of consumer learning (few repeat customers). In this framework, such a pattern leads us to infer there are differences in the degree of consumer learning across regions, before the grade cards.

A naive approach to implementing the test would be to regress the estimated values of α_j^A (ie. $\hat{\alpha}_j^A$) on a constant and the estimated values of α_j^B (ie. $\hat{\alpha}_j^B$). Deviations of $\hat{\alpha}_j^A$ from the fitted line may then indicate the presence of regional variation in reputational incentives, r_j. But deviations will also arise due to estimation error in the regional fixed effects ($\hat{\alpha}_j^B$ and $\hat{\alpha}_j^A$).

We therefore propose an approach that allows us to test for a linear relationship in the pre- and post-grade card fixed effects, that takes account of the estimation error in these fixed effects. Define RSS_u as the sum of squared residuals from the estimated equation (4) plus the sum of squared residuals from the estimated equation (5).[24] We then estimate a restricted specification, incorporating the restriction that α_j^A is a linear function of α_j^B:

$$\begin{aligned} s_{ijt} &= I_t^B \left(\alpha_j^B + \beta^B c_i + \gamma^B f_i + X_i\theta^B\right) \\ &+ \left(1 - I_t^B\right)\left(\kappa_1 + \kappa_2\alpha_j^B + \beta^A c_i + \gamma^A f_i + X_i\theta^A\right) + \epsilon_{ijt}, \end{aligned}$$

where I^B is an indicator for "before grade cards", and κ_1 and κ_2 are additional parameters to be estimated, in lieu of the post-grade card fixed effects. Equation (4) is nonlinear in the parameters, so we estimate this restricted specification via nonlinear least squares. Define RSS_r as the sum of squared residuals from the estimated equation (4).

Given our assumptions, if there is significant regional variation in reputational incentives, then an F-test will reject the hypothesis that RSS_u equals RSS_r. The advantage of this test over the naive approach, mentioned above, is incorporation of estimation error in the fixed effects (and all other parameters). The difference between this test and the above test with the assumption that $\alpha_3 = 0$, is that this test places no significance on absolute differences in region fixed effects. In this case, we focus on the relative impact of grade cards across regions. This is just another way of saying that if $\alpha_3 \neq 0$, then the prior test may lead us to incorrectly conclude there are regional differences in reputational incentives. Allowing for the possibility that $\alpha_3 \neq 0$

[24] Equivalently, we could combine equations (4) and (5) into a single equation, while allowing for different fixed effects and different coefficients on all variables before and after the grade cards. Then RSS_u is the same as the sum of squared residuals from this combined equation.

provides a more stringent test. The resulting F-statistic for this test of 5.77, leading us to reject the hypothesis with 99 percent confidence. This is not surprising, given the pattern shown in Figure 1. We conclude there is evidence of effective reputational incentives for some restaurants, due to the degree of consumer learning in the local region.

6.3 Regional Chain Effects

We can also apply the above tests of the regional effects to the chain restaurants. If reputational incentives are a significant determinant of restaurants' hygiene quality, and if the strength of these incentives vary across regions, then the incremental impact of reputational incentives due to chain affiliation may also vary across regions. Specifically, the impact of chain affiliation on hygiene quality, relative to non-chain restaurants, may be smaller in regions with a high degree of consumer learning (high r_j). In addition, in regions with a high degree of consumer learning, there may be less free-riding on chain reputation by franchisees.

In equation (1), the coefficient on the chain affiliation variable, β^B, measures the average difference in hygiene quality between chains and non-chains, while controlling for regional variation in average hygiene scores (α_j^B). In other words, β^B is the average within-region difference in hygiene scores between chains and non-chains. Similarly, the coefficient on the franchising variable, γ^B, measures the average within-region difference in hygiene scores between franchised chains and company-owned chains. In this section, we focus on the between-region variation in the chain and franchising effects. To do so, we estimate more general versions of equations (4) and (5), allowing for region-specific coefficients on the chain and franchising variables:

$$s_{ijt}^B = \alpha_j^B + \beta_j^B c_i + \gamma_j^B f_i + X_i\theta^B + \epsilon_{ijt}, \quad \text{and} \tag{8}$$

$$s_{ijt}^A = \alpha_j^A + \beta_j^A c_i + \gamma_j^A f_i + X_i\theta^A + \epsilon_{ijt}. \tag{9}$$

To identify the presence of regional differences in the degree of consumer learning, we assume the same functional forms as above. Specifically:

$$\begin{aligned}
\beta_j^B &= \beta_0 + \beta_1 r_j + \beta_2 w_j + \beta_3 r_j w_j, \\
\beta_j^A &= \beta_0 + \beta_1 g + \beta_2 w_j + \beta_3 g w_j, \\
\gamma_j^B &= \gamma_0 + \gamma_1 r_j + \gamma_2 w_j + \gamma_3 r_j w_j, \quad \text{and} \\
\gamma_j^A &= \gamma_0 + \gamma_1 g + \gamma_2 w_j + \gamma_3 g w_j.
\end{aligned}$$

We then implement the same set of tests that were applied to α_j^B and α_j^A, depending on whether we assume $\beta_3 = 0$ or $\gamma_3 = 0$. Note, the tests of the regional franchising effects are the most

informationally demanding of all the tests in this study. To identify this test, we rely on multiple observations of both company-owned and franchised chain restaurants in each region, for each of the major chains. It seems likely we will obtain imprecise estimates of these region-specific effects, which may undermine the proposed hypothesis tests.

Table 5 summarizes the results for all the tests in this study, including the tests discussed in this subsection. Under the assumption that $\beta_3 = 0$, we test whether $(\beta_j^A - \beta_j^B) \neq constant$. As shown in Table 5, the F-statistic is 1.66, allowing us to accept the inequality with 99 percent confidence. This indicates the reputational effect of chain affiliation varies across regions, according to the degree of consumer learning.

We can also check if the reputational effects embedded in the estimates of α_j^B are consistent with the reputational effects embedded in the estimates of β_j^B. Intuitively, in regions with a high degree of consumer learning, the incremental effect of chain affiliation on hygiene quality should be lower. To examine this, we assume that both $\alpha_3 = 0$ and $\beta_3 = 0$. This implies

$$\begin{aligned} \alpha_j^A - \alpha_j^B &= \alpha_1(g - r_j), \quad \text{and} \\ \beta_j^A - \beta_j^B &= \beta_1(g - r_j). \end{aligned}$$

Under these assumptions, if $(\alpha_j^A - \alpha_j^B)$ is negatively correlated with $(\beta_j^A - \beta_j^B)$, then α_1 and β_1 are of opposite signs. Presuming that α_1 is positive, we would infer the incremental effect of chain affiliation on hygiene quality is less in regions with high reputational incentives. Consistent with our interpretation, we compute the correlation coefficient to be -0.4163.

Allowing for the possibility that $\beta_3 \neq 0$, we also test if $\beta_j^A \neq \kappa_1 + \kappa_2 \beta_j^B$. The F-statistic is 0.36—we are unable to reject the hypothesis that $\beta_j^A = \kappa_1 + \kappa_2 \beta_j^B$. The test suggests there are not significant differences across regions in the effect of chain affiliation on reputational incentives.

Moving now to the regional franchising effects. First assuming that $\gamma_3 = 0$, so there is no interaction between consumer learning and willingness-to-pay in determining the extent of franchisee free-riding, we test if $(\gamma_j^A - \gamma_j^B) \neq constant$. The F-statistic, as reported in Table 5, is 0.52. Hence, we are unable to reject the hypothesis that franchising effects are the same across regions. On the one hand, this finding goes against the reputational incentives hypothesis. On the other hand, our analysis indicate the free-riding effect is relatively small in magnitude, so it is perhaps unsurprising that we do not find significant differences across regions.

We conclude that regional differences in the effect of chain affiliation on hygiene quality are supportive of the reputational incentives hypothesis. However, we are unable to detect regional differences in the franchise effect, casting a small degree of doubt on the effectiveness

of reputational incentives.

7 Conclusion

Prior empirical studies of reputational incentives focus on identifying the demand-side effect from reputation formation. In this study we present evidence in support of the supply-side effect of reputational incentives, using the example of restaurant hygiene quality. We perform a variety of tests aimed at distinguishing the reputational incentives hypothesis from alternative explanations for why some restaurants have good quality hygiene. The basis of our approach relies on the assumption that we, the econometricians, observe the very thing that consumers may only learn through the process of reputation formation. As we have argued, we believe this to be a reasonable assumption in this context.

We find that chain restaurants tend to have significantly better hygiene than independent restaurants due to the reputational effects of chain affiliation. This finding is robust to a number of alternative specifications, some of which utilize post-grade card hygiene scores to control for unobserved heterogeneity. To the extent that chain reputation is a source of competitive advantage for chain relative to non-chain restaurants, the introduction of posted grade cards reduces this advantage for chains. Prior to the grade cards, company-owned chain units have slightly better average hygiene than franchised chain units. Most of our analysis indicates this is due to franchisee free-riding on chain reputation, although in some specifications the effect is insignificant. These findings provide additional verification of the reputational effect of chain affiliation.

To identify a possible effect of reputational incentives on hygiene quality at independent restaurants, we assume the degree of consumer learning is a characteristic of the local region, defined as either a city or a zip code. We also assume the grade cards equalize consumer information across regions, eliminating any differences in restaurants' reputations derived from pre-grade card consumer learning. Our analysis indicates there are significant differences across regions in the degree reputation formation for restaurants. We also show that regions where independent restaurants tend to have relatively good quality hygiene, the incremental effect on hygiene from chain affiliation is lower. This suggests the reputational effect of chain affiliation is reduced in regions with a high degree of consumer learning. These findings support our conclusion that reputational incentives also apply to independent restaurants.

While there may be features of the Los Angeles restaurant industry, before grade cards, that are idiosyncratic, it seems plausible to us that these findings can be extrapolated to restaurants

in other regions where consumers have limited information about hygiene inspection results. More generally, reputation mechanisms are thought to be important in numerous markets where consumers are uninformed about product quality. Our findings support the view that reputations can cause firms to provide good quality products. However the fact remains, the introduction of restaurant hygiene grade cards in Los Angeles in 1998 had a large positive effect on hygiene quality (Jin and Leslie, 2003). Without grade cards, many restaurants had poor hygiene. One explanation may be that reputations are effective at causing good hygiene for restaurants that face a high degree of consumer learning, but most restaurants are not in this situation. Hence, while reputational incentives are effective, posting of DHS hygiene grade cards solves the problem that most restaurants are unable to form a reputation for good hygiene if they were to provide it.

Bibliography

Bajari, P. and A. Hortacsu (2003): "Online Auctions: A Review of Empirical Findings," forthcoming in *Journal of Economic Literature*.

Bond's Franchise Guide, 2001, 13th version, Source Book Publications.

Borenstein, S. and M.B. Zimmerman (1988): "Market Incentives for Safe Commercial Airline Operation," *American Economic Review*, 78(5), 913–35.

Brickley, J.A. and F.H. Dark (1987): "The Choice of Organizational Form: The Case of Franchising," *Journal of Financial Economics*, 18, 401–20.

Cabral, L. and A. Hortacsu (2004): "The Dynamics of Seller Reputation: Theory and Evidence from eBay," *Mimeo*.

Cai and Obara (2004)

Caves, R.E. and W.F. Murphy (1976): "Franchising: Firms, Markets, and Intangible Assets," *Southern Economic Journal*, 42(4), 572–586.

Chernew, M., G. Gowrisankaran and D.P. Scanlon (2002): "Learning and the Value of Information: The Case of Health Plan Report Cards," *Mimeo*.

Dewally, M. and L. Ederington (2002): "A Comparison of Reputation, Certification, Warranties, and Disclosure as Remedies for Information Asymmetries: Lessons from the Online Comic Book Market," *Mimeo*.

Hubbard, T. (2002): "How Do Consumers Motivate Experts? Reputational Incentives in an Auto Repair Market," *Journal of Law and Economics*, October, 437–68.

Israel, M. (2003): "Consumer Learning about Established Firms: Evidence from Automobile Insurance," *Mimeo*, Kellogg School of Management.

Jin, G. and P. Leslie (2003): "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, 118(2), 409–51.

Klein, B. and K. Leffler (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89(4), 615–641.

Lafontaine, F. (1992): "Agency Theory and Franchising: Some Empirical Results," *RAND Journal of Economics,* 23(2), 263–283.

Lafontaine, F. and K.L. Shaw (1999): "The Dynamics of Franchise Contracting: Evidence from Panel Data," *Journal of Political Economy*, 107(5), 1041–1080.

Lafontaine, F. and K.L. Shaw (2003): "Targeting Managerial Control: Evidence from Franchising," forthcoming in *RAND Journal of Economics*.

Mailath, G.J. and L. Samuelson (2001): "Who Wants a Good Reputation," *Review of Economic Studies*, 68(2), 415–441.

Norton, S.W. (1988): "An Empirical Look at Franchising as an Organizational Form," *Journal of Business*, 61(2), 197–218.

Resnick, P. and R. Zeckhauser (2002): "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System," in M.R. Bayes (ed), *The Economics of the Internet and E-Commerce*, Elservier Science, Amsterdam, 11:127–157.

Rubin, P. (1978): "The Theory of the Firm and the Structure of the Franchise Contract," *Journal of Law and Economics*, 21(1), 223–33.

Stigler, G.J. (1961): "The Economics of Information," *Journal of Political Economy*, 69, 213–25.

Tadelis, S (1999): "What's in a Name? Reputation as a Tradeable Asset," *American Economic Review*, 89(3), 548–63.

Table 1: Summary of Hygiene Scores, 1995 to 1998

	Number of restaurants	Number of inspections	Mean (std dev) score before grade cards		Mean (std dev) score after grade cards	
All restaurants	24,304	127,111	76.77	(14.72)	89.62	(7.68)
Zagat restaurants	1,017	4,493	77.43	(14.10)	88.97	(7.54)
All chain (cmpny-ownd)	1,655	9,797	82.94	(11.53)	92.70	(5.65)
All chain (franchised)	977	5,635	81.84	(12.67)	92.87	(5.58)
Brgr King (cmpny-ownd)	64	353	86.98	(9.59)	94.04	(4.06)
Brgr King (franchised)	61	389	82.09	(11.29)	94.14	(4.38)
El P Loco (cmpny-ownd)	95	614	82.73	(11.35)	93.15	(4.34)
El P Loco (franchised)	24	175	77.82	(13.82)	92.17	(4.58)
Jack in Box (cmpny-ownd)	109	669	83.63	(11.96)	94.82	(3.68)
Jack in Box (franchised)	39	229	82.10	(12.43)	93.21	(5.38)
KFC (cmpny-ownd)	85	566	81.49	(11.43)	90.83	(6.65)
KFC (franchised)	49	314	78.12	(13.75)	92.04	(5.60)
McDonalds (cmpny-ownd)	109	746	81.09	(12.16)	91.50	(5.91)
McDonalds (franchised)	147	883	81.78	(11.78)	92.69	(5.22)
Taco Bell (cmpny-ownd)	131	736	85.39	(10.49)	95.25	(4.04)
Taco Bell (franchised)	42	238	85.44	(12.04)	95.58	(4.00)
Burger restaurants	1,283	7,982	78.77	(14.13)	91.30	(5.32)
Chicken restaurants	320	2,014	78.94	(12.78)	90.67	(6.26)
Chinese restaurants	818	5,449	70.68	(16.51)	86.13	(8.78)
Mexican restaurants	1,592	9,752	74.83	(15.19)	88.92	(8.08)
Pizza restaurants	1,098	6,452	79.26	(12.83)	90.87	(6.54)
0–30 seats	13,019	66,271	77.43	(14.39)	90.06	(7.54)
31–60 seats	5,444	29,714	75.61	(14.97)	89.05	(7.77)
61+ seats	5,841	31,126	76.46	(15.13)	89.29	(7.80)
Lower income areas	12,130	60,993	74.55	(15.30)	89.78	(7.79)
Higher income area	12,174	66,118	78.79	(13.87)	89.47	(7.56)

Table 2: Variance Decomposition of Pre-Grade Card Hygiene Scores

	Number of Variables	Sum of Squared Residuals	R^2
All Pre-Grade Card Inspections:			
Constant	10	17,402,286	0.0419
Restaurant characteristics	46	16,155,534	0.1106
City fixed effects	161	14,614,335	0.1954
Zip code fixed effects	325	13,298,310	0.2679
Restaurant fixed effects	22,211	6,826,502	0.6242
Number of observations	83,790		
Regime I — Objective and Subjective Components:			
Constant	7	15,274,943	0.0006
Restaurant characteristics	42	14,205,550	0.0706
City fixed effects	157	12,293,296	0.1957
Zip code fixed effects	314	10,885,802	0.2878
Restaurant fixed effects	21,030	4,799,013	0.6860
Number of observations	66,977		
Regime I — Objective Component Only:			
Constant	7	4,337,273	0.0010
Restaurant characteristics	42	4,079,308	0.0605
City fixed effects	157	3,312,982	0.2370
Zip code fixed effects	314	2,799,496	0.3552
Restaurant fixed effects	21,030	1,264,006	0.7089
Number of observations	66,977		

All specifications also include a full set of year-qtr dummies. The specifications in the top panel also include inspection regime dummies.

Table 3: Predictors of Pre-Grade Card Hygiene Scores

	Estimated coefficient	Standard error
Constant	85.3395	3.1635***
Regime II dummy	9.8947	0.2238***
Chain restaurant	4.2756	0.3356***
Franchised chain	-0.6439	0.2825**
Zagats restaurant	1.8108	0.9459**
Zagats food rating	-0.0839	0.0500*
Licensed to sell beer	-3.4148	0.2552***
Licensed to sell liquor	-2.4120	0.2716***
American-style restaurant	0.7307	0.3378**
Burger restaurant	0.1486	0.2619
Chinese restaurant	-3.4536	0.2614***
Coffee restaurant	2.0558	0.4148***
Mexican restaurant	-0.1192	0.2253
Pizza restaurant	0.9928	0.2645***
0–30 seats	1.4734	0.1204***
61+ seats	-.3541	0.1391**
Fraction of zip employment in hotels	31.6130	2.3746***
Fraction of zip employment in museums	-3.1671	14.1231
Fraction of zip employment in amuseprks	19.1298	10.3208*
Mostly chain restaurants in zip	1.7873	0.1217***
Mostly independent restaurants in zip	-3.2265	0.1632***
Mostly company-owned chains in zip	1.4620	0.1071***
Mostly franchised chains in zip	-2.5329	0.1241***
Mean household size in census tract	-3.7954	0.1733***
Mean personal income in tract	1.32e-5	7.99e-6*
Fraction married in census tract	18.3543	0.6717***
Fraction age 65+ in census tract	-28.2514	1.0571***
Fraction Asian in census tract	-9.1733	0.4466***
Fraction Hispanic in census tract	-5.0647	0.4199***
Number of observations	83,691	
Adjusted R^2	.2015	

Unreported variables: year-quarter dummies, chicken restaurants, Latin food, deli-style restaurants, family-style restaurants, bakery, missing food type dummy, and home delivery dummy.

Stars denote significance levels: 99 percent confidence level (***), 95 percent confidence level (**) and 90 percent confidence level (*).

Table 4: Determinants of Restaurant Hygiene Scores

	Coefficient	Std error	Coefficient	Std error	Coefficient	Std error	Coefficient	Std error
Chain restaurant	4.8834	.3275***			5.3894	.8811***	3.8234	.3243***
Franchised chain restaurant	-.5730	.2766**	-2.3162	.6772***	-1.7100	.4570***	-0.1636	.2714
Grade cards								
Grade cards × chain			-4.0567	.4645***				
Grade cards × franchised			1.2332	.3881***				
Mean post-grade card score					.4922	.0496***	.4868	.0078***
Number of observations	83,790		127,111		77,255		77,255	
R^2	.2435		.6021		.1546		.2871	
City fixed effects	Yes		Absorbed		No		Yes	
Restaurant fixed effects	No		Yes		No		No	
Restaurant characteristics	Yes		Yes		Yes		Yes	
Grade cards × rest. chars	No		Yes		No		No	
Pre-grade card observations	Yes		Yes		Yes		Yes	
Post-grade cards observations	No		Yes		No		No	

All specifications include year-quarterly dummies. The *grade cards* variable drops out of the second specification due to collinearity with the time dummies.

In the third and fourth specifications, post-grade card observations enter in the construction of the independent variable *Mean post-grade card score*. Only pre-grade card scores are used in the dependent variable.

Stars denote significance levels: 99 percent confidence level (***), 95 percent confidence level (**) and 90 percent confidence level (*).

Table 5: Summary of Hypothesis Tests

Reputation prediction	Relevant equations	Test statistic	
$\beta^B > 0$	(1)	14.91***	(t-test)
$\gamma^B < 0$	(1)	-2.07**	(t-test)
$\beta_2 < 0$	(2)	-8.83***	(t-test)
$\gamma > 0$	(2)	3.18***	(t-test)
$\beta > 0$	(3)	11.79***	(t-test)
$\gamma < 0$	(3)	0.60	(t-test)
$(\alpha_j^A - \alpha_j^B) \neq constant$	(4) to (7)	38.00***	(F-test)
$\alpha_j^A \neq \kappa_1 + \kappa_2 \alpha_j^B$	(4) to (7)	5.77***	(F-test)
$(\beta_j^A - \beta_j^B) \neq constant$	(8) & (9)	1.66***	(F-test)
$\beta_j^A \neq \kappa_1 + \kappa_2 \beta_j^B$	(8) & (9)	0.36	(F-test)
$(\gamma_j^A - \gamma_j^B) \neq constant$	(8) & (9)	0.52	(F-test)

Stars denote significance levels: 99 percent confidence level (***), 95 percent confidence level (**) and 90 percent confidence level (*).

Figure 1: Mean Hygiene Scores for each City
in Different Time Periods