# Valuing the Consumption Benefits of Urban Density

**Victor Couture** *

*University of Toronto*

JOB MARKET PAPER, preliminary

This draft: January 2013

ABSTRACT: Density is a defining feature of cities, yet there is little evidence as to how consumers benefit from urban density. This paper investigates the consumption value of density by combining travel data with local business data from Google Places. I first show that increased density enables consumers to both realize gains from variety and save time through shorter trips. I then estimate the gains from density in the restaurant industry, identifying the value of access to a preferred location from an individual's willingness to incur extra travel costs to reach it. The results reveal wide disparities across areas in a variety-adjusted restaurant price index, leading to significant geographic welfare differentials. Within large metropolitan areas, the price index generally drops by more than 20% from a city's periphery to the denser downtown core. This decrease represents yearly gains of about $400 for an average household, considering restaurants only. The model predicts a key feature of the data, that increasing the density of destinations generates little reduction in trip times. Most of the gains from density are therefore gains from variety, not savings on travel time. Americans' aggregate welfare gains from access to a variety of eating options beyond the one restaurant closest to them amount to approximately 2% of consumer expenditures, the first estimate of the gains from variety in the service sector.

Key words: consumer cities, gains from variety, urban density, accessibility, travel demand.

JEL classification: D12, R41

*I am grateful to my advisors Gilles Duranton, Robert McMillan, Matthew Turner and Will Strange for their advice and encouragement throughout this project. The paper also benefited from conversations with Victor Aguirregabiria, Nate Baum-Snow, Ettore Damiano, Vernon Henderson, Maciej Kotowski, Nicholas Li, Martin Osborne, Diego Puga, Olmo Silva, Aloysius Siow and Nathan Yang. I also thank seminar participants at Brown University and the University of Toronto for additional comments. Financial support from the Social Science and Humanities Research Council and from the Royal Bank Graduate Fellowship in Public and Economic Policy is gratefully acknowledged. All remaining errors are my own.

Contact: Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario M5S 3G7, Canada (e-mail: victor.couture@utoronto.ca)

## 1. Introduction

Density is a fundamental characteristic of cities. In dense urban areas, workers are more productive, and a sophisticated literature relates much of this productivity premium to the advantages of spatial proximity.[1] Higher density may also improve consumers' access to a variety of goods and services (Glaeser, Kolko, and Saiz, 2001), but despite vigorous debate over the potential for consumption amenities to drive urban success,[2] direct evidence is scarce as to the origin and importance of the consumption advantage of cities.

This paper sets out a new approach to estimating the consumption value of urban density. The estimation uses travel data and exploits the recent availability of detailed online microgeographic data on local businesses. While it is simple to compute the value of shorter trip times, my methodology also allows me to estimate the gains from increased choice in denser areas (so-called 'gains from variety'). I identify an individual's willingness to pay for access to a preferred location from the extra travel costs that she incurs to reach it. I provide estimates of the gains from density in the US restaurant industry, a prominent part of the urban service sector.

This exercise is of interest for several reasons. American households spend a significant share of their income on non-tradable services such as restaurants, live entertainment, and many professional services (e.g. medical care) requiring face-to-face interactions, with restaurants alone accounting for more than 5% of household expenditures. Non-tradables presumably represent most of the consumption value of modern cities, given the dramatic decline in the cost of shipping goods over the last century. My analysis highlights how urban density facilitates the movement of people, on which much of the service sector depends.

Research on the consumption benefits of density also has important implications for our understanding of travel behavior. The idea that urban density allows individuals to cheaply substitute among travel destinations explains a striking feature of the data, namely that increasing the density of available destinations fails to induce sizable reductions in trip times.[3] Spatial proximity to restaurants in urban areas could allow individuals to make very short trips to eat

[1]See for instance Melo, Graham, and Noland (2009) for a meta-analysis of estimates of agglomeration economies, and Combes, Duranton, and Gobillon (2011) for a survey of key empirical issues.

[2]For instance, Clark (2003) shows that cities providing more natural and constructed amenities experience faster population growth. Carlino and Saiz. (2008) find that metropolitan areas that are attractive to tourists (likely because of consumer amenities) are also growing faster.

[3]I define density in terms of restaurants per minute of travel, not restaurants per mile. For instance, holding all else constant a decrease in car travel speed reduces density.

out, but I find that they often choose to benefit from higher restaurant density by visiting a more distant location that they prefer. That is, the welfare gains from density are mostly gains from variety, as opposed to travel time savings on each trip. This distinction is relevant to the evaluation of policy proposals designed to reduce vehicle travel by encouraging higher density living.

Finally, my results on the gains from variety in the service sector contribute to an emerging literature estimating the gains from variety in consumer goods (Broda and Weinstein, 2006, 2010), and provide empirical support for the widespread use of utility functions featuring a preference for variety.[4]

To estimate the gains from density, I specify a discrete-choice model of demand for restaurants. In the model, each restaurant receives a logit utility shock, and locations farther away from an individual are more expensive because of travel costs. Individuals face a trade-off between the gains from visiting a preferred restaurant and the costs of a long trip. The key parameter of the model is an elasticity of substitution between restaurants, which I estimate by maximum likelihood. This estimation does not require data on restaurant choice, only on trip length. If individuals always travel to the closest restaurant, then restaurants must be perfect substitutes and there are no gains from restaurant variety in dense areas, only savings through shorter trips. If individuals take long trips to eat out, then restaurants are imperfect substitutes and gains from variety are correspondingly large.[5]

Estimating a discrete-choice model for travel destinations requires not only data on travel behavior, but also comprehensive microgeographic data on the location of all destinations available to an individual. The travel data is from the National Household Travel Survey (NHTS), which identifies trips to a restaurant and the location of an individual at the block group level. The NHTS also allows the estimation car travel speed in different areas. I collect data on each restaurant online, from its Google Places page.[6] I observe the exact location of almost all restaurants (273,000 units) in 15 states representing 50% of the US population.[7] For robustness tests and

[4]Preferences for variety are often modeled after Dixit and Stiglitz (1977). See Broda and Weinstein (2006) for an empirical study of the gains from import variety, and Broda and Weinstein (2010) for an empirical study of the gains from consumer goods variety.

[5]An endogeneity problem typically arises when estimating an elasticity of substitution, due to the unobserved relationship between higher prices and better quality. In general, one should not interpret individuals' insensitivity to price differentials as a preference for variety, but rather as a willingness to pay for quality. With travel data however, some variation in restaurant prices comes from transport costs that are plausibly unrelated to quality differentials.

[6]As of May 2012, Google's local business pages are called Google+ Local pages.

[7]I choose these states because a vast majority of the trips in my travel data originate there. See section 3.

extensions of the model, I also require data on restaurant characteristics. Each restaurant's name and category (e.g. 'Pizza') comes from its Google page. For a subset of my sample, there is additional information such as meal price and quality ratings on Yelp, a popular review website.

To obtain estimates of the consumption value of urban density, I derive a variety-adjusted restaurant price index from the logit model, which turns out to be identical to the 'love-of-variety' constant elasticity of substitution (CES) price aggregator.[8] The restaurant price index in a location is low if there are many restaurants nearby. Higher travel speed also decreases the price index. I find large variations across areas in the price index, generating substantial spatial welfare differentials. The gains from density are very localized, and much of the variation in the price index occurs within large metropolitan areas (MSA). For a car driver, the price index generally drops by more than 20% from a city's outskirts to its downtown. Considering restaurants only, such density increase represents yearly gains between $300 and $500 for an average household. These findings are consistent with Albouy and Lue's (2011) quality of life estimates, which are higher in denser areas and vary almost as much within metropolitan areas as across them.[9]

There are few existing estimates with which to compare my results on the consumption benefits of density. A large literature measures the productivity advantages of cities, but their consumption advantages has received less academic attention. Handbury and Weinstein (2012) find that residents of larger cities face a lower price index for consumer goods (groceries), controlling for store amenities, individual characteristics and differences in the number of varieties available. Their price index drops by 10% from New York to Des Moines, the smallest MSA in their sample. They estimate that New York residents have access to 97,000 types of groceries within the MSA, versus 32,000 in Des Moines, leading to a 3% decrease in the price index that is purely due to variety. While these numbers are estimated for large areas and do not take transport costs into account, simple comparisons with my results for the restaurant industry hint at much larger geographic welfare differentials in the non-tradable service sector. Residents of rural areas sometimes face a restaurant price index 100% higher than that in the densest urban areas, in part because of hefty transport costs to reach even the closest locations. Many residents of the New York or Los Angeles metropolitan areas have access to more than 10,000 restaurants within 45

[8] Anderson, de Palma, and Thisse (1992) prove that the under a linear utility specification - that I use - the logit and CES model lead to the same choice probabilities.

[9] Albouy and Lue (2011) obtain quality of life estimates using wage, housing-cost and commuting-cost differentials. Regressions on these quality of life estimates also suggest that individuals are willing to pay for access to a high density of bars and restaurants.

minutes of car travel, and would face at least a 25% reduction in the restaurant price index from moving anywhere in Des Moines, mostly because of a decrease in variety. Des Moines offers faster car travel, but still only 650 restaurants within 45 minutes of travel. These results support the argument in Glaeser *et al.* (2001) that cities have an edge in the service sector because of lower transport costs in denser areas.

To obtain an aggregate measure of the welfare gains from restaurant variety, I compute how much Americans are willing to pay each year to access additional restaurants beyond the closest eating option available. I find values ranging from $80 to $160 billion. Realizing these (net) benefits requires 5.6 billion hours of extra travel past the nearest restaurant, worth $67 billion at a value of time of $12 per hour. Such non-work trips represent 70% of all travel in the US, and the overwhelming majority of these trips are by car. Private vehicles are responsible for roughly 20% of national greenhouse gas emissions from fossil fuels, as well as externalities in the form of pollution, accidents and congestion.[10] From a policy-making perspective, my results draw attention to a major obstacle facing any attempt at curbing vehicle travel: travel costs are low relative to the gains from variety, and especially so in urban areas.

The logit model's estimates of the gains from density implicitly assume that individuals benefit from density in part by visiting preferred destinations. If the model is correct, then increasing restaurant density should induce little reduction in travel times. The data allow me to test this hypothesis. The model predicts the probability of a trip to each restaurant in each area, for a given estimate of the elasticity of substitution between restaurants. Knowing the predicted probability of a trip of each length for each traveler in my sample, I can obtain a set of predicted trip times for comparison with actual trip times. To test the model, I run OLS regressions of predicted trip time on measures of restaurant density, and compare the coefficients with those from the same regressions on actual data.[11] This exercise identifies a discrepancy between the data and the model's prediction, in that individuals living in denser areas make somewhat shorter trips than

[10]See Parry, Walls, Winston Harrington, and Policies (2007) for a discussion of car externalities. The data on greenhouse gases is from the US Environmental Protection Agency's Inventory of US Greenhouse Gas Emissions and Sinks: 1990-2010, which is available for download at `http://www.epa.gov/climatechange/emissions/usinventoryreport.html`. In this paper, I find that for most Americans, fast private vehicles are necessary to realize sizable gains from restaurant variety, but this situation is not inevitable. A walker in Manhattan, the highest density area, faces a lower price index than a driver in 95% of other areas of the country.

[11]In general, one cannot fit a model using a given dataset, and then test it using the same data. In this case however, regressions on predicted trip time produce coefficients that are almost invariant (except for the constant term) to the choice of parameters used to obtain these predicted trip times.

what the model predicts.

I therefore extend the model in several directions to better match the data and to investigate the possibility of omitted variables. The starting point for the first extension is a comparison of OLS and IV regression results on the determinants of trip time. The instrument for restaurant density is past growth in population density, for a selected sample of travelers with a very low probability of moving. In instrumented regressions, the effect of density on trip time is more positive and much closer to the model's prediction, suggesting that individuals with a high value of travel time, who make shorter trips, sort into dense areas. Based on these results, I estimate a version of the logit model with sorting by value of travel time into areas with different restaurant density. The next two extensions allow remote restaurants to be close substitutes to similar restaurants (e.g. serving the same type of cuisine) nearer to home, which could also explain why individuals in dense areas make shorter trips than what the model predicts. I first estimate a model in which restaurants in the same chain are perfectly substitutable. Then I specify and estimate a nested-logit model in which restaurants within the same category (e.g. 'Sushi') are more substitutable. I infer exogenous tastes for categories from a free-entry equilibrium condition for restaurant suppliers. Finally, I use partial price data from Yelp to let meal prices vary with restaurant density, or with travel time from an individual to a restaurant. These extensions ultimately lead to similar welfare results, except for the nested-logit model whose parameters are hard to estimate precisely with my data.

These results on the consumption benefits of density are relevant to two major strands of literature in urban planning and transportation. First, one can interpret my variety-adjusted price index for destinations as what transportation researchers call a 'travel accessibility index,' with many desirable characteristics. The variety-adjusted price index accounts for differences in car travel speed across areas, and places less weight on remote destinations. Most important, the variety-adjusted price index has a microeconomic foundation. Unlike available travel accessibility indices, it has a natural interpretation as a price, and it depends on the structural parameters of a model.[12] Second, reduced-form regressions of trip time on measures of restaurant density belong to a vast empirical literature measuring the relationship between vehicle mile traveled (VMT) and the built environment. The regressions in this paper differ from those in extant research in that they use the exact location of destinations to measure density, instead of proxies like population

[12] See Bhat, Handy, Kockelman, Mahmassani, Chen, and Weston (2000) for a review of the literature on travel accessibility indices, and for many examples of such indices proposed in transportation research.

and job density. Consistent with my results, other studies find a very small effect of density on VMT, and I show that a model of travel demand with substitutable destinations provides a theoretical framework to understand this empirical regularity.[13]

## 2. A logit model of travel demand

My analysis starts from two assumptions about demand for restaurant travel. The first is that restaurants are substitutable, so that an individual prefers some restaurants to others. The second is that travel is costly, so that the price of visiting a restaurant farther away is higher. Irrespective of the details of the model, these assumptions imply a trade-off between the gains from going to a preferred restaurant and the costs of a long drive. At higher density, the price difference between two adjacent restaurants is lower. This means that travelers in dense areas have a lower probability of visiting the restaurant closest to home, which has the lowest transport costs but is not necessarily preferred. Increasing density induces little reduction in trip time, but generates gains from variety, i.e. from visiting preferred location.

Consider an individual living in area $k$ and choosing a restaurant.[14] Let $i$ index the number $I_k$ of restaurants available, so that $i \in \{1,2,3,\ldots,I_k\}$. The restaurant with index $i$ = 1 is closest, $i$ = 2 is second closest and so on. Denote travel time to restaurant $i$ by $t_{ki}$, and the price of a meal at any restaurant by a constant $p$.[15] The total price of eating at restaurant $i$, including transport costs to and from the restaurant, is $p_{ki} = p + 2\gamma t_{ki}$, where $\gamma$ is the value of travel time. This total price is what should be understood when I refer to restaurant price elsewhere in the paper, unless I mention 'meal' price specifically.

Each restaurant receives a random idiosyncratic shock $\epsilon_{ki}$, which captures an individual's preference for restaurant $i$. $\epsilon_{ki}$ is a random draw from a type I extreme value distribution with scale parameter $1/(\sigma - 1)$, where $\sigma$ will turn out to be the elasticity of substitution between

[13]See Ewing and Cervero (2010) for a meta-analysis. They find a weighted average elasticity of VMT of -0.04 with respect to household/population density and 0.00 with respect to job density (the area over which density is computed varies in each study). The literature also finds that the number of and purpose of trips that individual undertake vary little across areas.

[14]In Appendix B, I extend the model to cover the decision of how much to spend on restaurants and how much to spend on all other goods. This extension allows me to provide welfare estimates that account for the degree of substitution between restaurants and all other goods.

[15]Assuming a constant meal price is equivalent to assuming that utility is invariant to variation in meal prices, because higher (lower) prices are always exactly compensated by higher (lower) quality. Such an assumption is reasonable if quality is mostly produced through variable costs, as is likely the case in the restaurant industry (see Berry and Waldfogel (2010) for a discussion of product quality in the restaurant industry).

restaurants. Following Anderson *et al.* (1992), define the utility from making $r_{ki}$ trips to restaurant $i$ as:

$$u_{ki} = ln(r_{ki}) + \epsilon_{ki}.$$

Let $y_k$ be expenditures on restaurants for an individual in area $k$, so that $p_{ki}r_{ki} = y_k$ is the individual's budget constraint. Substituting $r_{ki}$ from the budget constraint into the utility function leads to the following indirect utility from choosing restaurant $i$:

$$v_{ki} = ln(y_k) - ln(p_{ki}) + \epsilon_{ki}.$$

This individual's utility maximization problem is:

$$max\{-ln(p_{k1}) + \epsilon_{k1}, \ldots, -ln(p_{ki}) + \epsilon_{ki}, \ldots ln(p_{kI_k}) + \epsilon_{I_k}\}. \tag{1}$$

Note that expenditures $y_k$ are constant within an area and do not affect restaurant choice. The logit choice probability is equal to $\frac{e^{ln(p_{ki})/(\sigma-1)}}{\sum_{i=1}^{I_k} e^{ln(p_{ki})/(\sigma-1)}}$ for all restaurants $i$ (see Train (2009) for details and a proof). Restaurant $i$ is chosen with probability $\frac{p_{ki}^{1-\sigma}}{\sum_{i=1}^{I_k} p_{ki}^{1-\sigma}}$ and the number of trips is equal to $r_{ki} = y_k/p_{ki}$, so the probability of a trip to restaurant $i$ of length $t_{ki}$ in area $k$, given the set of travel times to all restaurants: $T_k = \{t_{k1}, \ldots, t_{ki}, \ldots, t_{kI_k}\}$ is:

$$prob_{ki} = prob(t_{ki}|T_k) = \frac{p_{ki}^{-\sigma}}{\sum_{i=1}^{I_k} p_{ki}^{-\sigma}}. \tag{2}$$

The probability ratio of a trip to two restaurants $i$ and $j$ is:

$$\frac{prob_{ki}}{prob_{kj}} = \left(\frac{p_{ki}}{p_{kj}}\right)^{-\sigma} = \left(\frac{p + 2\gamma t_{ki}}{p + 2\gamma t_{kj}}\right)^{-\sigma}. \tag{3}$$

As an example, if $i$ = 1 and $j$ = 2, then $\frac{prob_{k1}}{prob_{k2}} = \left(\frac{p+2\gamma t_{k1}}{p+2\gamma t_{k2}}\right)^{-\sigma}$ is the ratio of the probability that an individual travels to the closest restaurant to the probability that she travels to the second closest. If $\sigma = \infty$, then restaurants are perfect substitutes and the ratio in equation (3) tends to infinity as the individual travels exclusively to the closest restaurant. If the difference between $t_{k1}$ and $t_{k2}$ is large, in a low-density area in which restaurants are far apart, then the proportion of trips to restaurant 1 will be larger. That is, in low-density areas individuals mostly visit the closest restaurants, because substituting between restaurants is expensive.

It is easy to use equation (3) to show that $\sigma$ represents the elasticity of substitution between restaurants. In the model, an individual has constant tastes and always travels to the same restaurant, so a low $\sigma$ represents heterogeneous preferences across many individuals. However, one can equivalently think of one individual getting new idiosyncratic shocks from the same distribution before each restaurant choice, in which case a low $\sigma$ also represents a taste for variety.[16] The price index that I derive next does not distinguish between these two interpretations, and neither do my empirical results.

### *A price index*

The indirect utility function of an individual is equal to the expected value of the indirect utility $v_{ki}$, given that individuals choose the restaurant that maximizes equation (1):

$$\mathbf{E}\left(\max_{i\in\{1,\ldots,I_k\}}\{v_{ki}\}\right) = ln\left(y_k\left(\sum_{i=1}^{I_k} p_{ki}^{1-\sigma}\right)^{(1/(1-\sigma))}\right). \tag{4}$$

Suppose that there are two areas, $k$ and $k'$, each with its own restaurant choice set, of size $I_k$ and $I_{k'}$. Define a relative price index $P_{k,k'}$ as the factor by which the restaurant prices in area $k$ would have to change in order to equalize indirect utility in both areas, assuming constant expenditures on restaurant. That is:

$$ln\left(y\left(\sum_{i=1}^{I_k}(P_{k,k'}p_{ki})^{1-\sigma}\right)^{(1/(1-\sigma))}\right) = ln\left(y\left(\sum_{i=1}^{I_{k'}} p_{k'i}^{1-\sigma}\right)^{(1/(1-\sigma))}\right),$$

so that:

$$P_{k,k'} = \frac{\left(\sum_{i=1}^{I_{k'}} p_{k'i}^{1-\sigma}\right)^{1/(1-\sigma)}}{\left(\sum_{i=1}^{I_k} p_{ki}^{1-\sigma}\right)^{1/(1-\sigma)}}. \tag{5}$$

Note that $P_{k,k'}$ is exactly the relative price index that would be derived from CES preferences.[17] To see how this index captures welfare gains from restaurant density, suppose that there is an infinity of restaurants equally spaced on a line, 2 minutes from one another in area $k$ and 4 minutes from one another in area $k'$, so that the density of restaurants is higher in area $k$. A restaurant with any given index $i$ has a lower price in the denser area $k$, because $p_{ki} = p + 4\gamma i > p_{k'i} = p + 8\gamma i$. This is true for all $i$, so the denominator of equation (5) must be smaller than its numerator, and $P_{k,k'}$ is smaller than 1.

[16]The logit model can be thought of as a microfoundation for CES preferences. Dixit and Stiglitz (1977), in their original paper, also mention that CES preferences can equivalently represent an aggregation of many individuals' heterogeneous preferences, or the preferences of one individual with a taste for variety.

[17]I do not use the exact price index proposed by (Sato (1976) and Vartia (1976)) for CES preferences, or the equivalent index that is robust to the introduction of new goods, introduced by Feenstra (1994). These indices are useful because they provide expressions in which the expenditure shares on each variety capture an unobserved quality parameter, which disappears from the expression of the exact price index. In my framework, however, the particular quality parameter of any given restaurant is less relevant, because variation in prices come from variation in travel time from an individual, not from unobserved quality. I include such a quality parameter in the nested-logit model of section 7.

It is useful to define the numerator in equation (5) as a price index in area $k$, denoted by $P_k$, and the denominator as a variety-adjusted price index in area $k'$, denoted by $P_{k'}$, so that for instance:

$$P_k = \left( \sum_{i=1}^{I_k} (p + 2\gamma t_{ki})^{1-\sigma} \right)^{1/(1-\sigma)} . \tag{6}$$

Note that this price index has another interpretation as what transportation researchers call a 'travel accessibility index',[18] with a number of desirable characteristics. First, the total price of reaching a destination depends on travel time, not just on distance, so the price index allows for speed differences across areas. Second, the impact of a destination on accessibility decreases with how far it is from home, and one can estimate the strength of this decay. If $\sigma$ is high, then only destinations very close to home affect $P_k$, and the contribution of remote destinations is negligible. If $\sigma$ is very low then even distant destinations contribute to accessibility. Third, $P_k$ has a micro-foundation. Both $\sigma$ and $\gamma$ are structural parameters with an economic interpretation, $\gamma$ as a monetary value of travel time, and $\sigma$ as an elasticity of substitution. $P_k$ itself is an economically meaningful quantity (a price), that is sometimes called the 'unit cost of services' when derived from CES preferences.

***Theoretical predictions of the determinants of trip time***

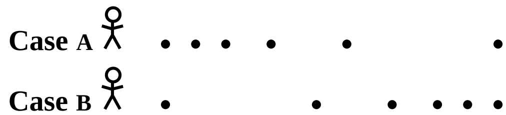
A model in which density allows individuals to realize gains from variety has strong implications for the effect of density on trip time. To develop an understanding of the model's predictions on the relationship between expected trip time and the restaurant choice set, I now prove theoretical results that hold exactly for simple distributions of restaurants (e.g. uniform). In section 6, I obtain the model's predictions for the true distribution of restaurants in the data, that I use to test the model.

The expected length of a trip to a restaurant is the sum of the probability of each trip multiplied by its travel time, so average trip time in area $k$ is equal to:

$$\bar{t}_k = \frac{\sum_{i=1}^{I_k} (p + 2\gamma t_{ki})^{-\sigma} t_{ki}}{\sum_{i=1}^{I_k} (p + 2\gamma t_{ki})^{-\sigma}}. \tag{7}$$

[18]Ben-Akiva and Lerman (1985) propose to use denominator of a logit probability in a travel model as a travel accessibility index, and Niemeier (1997) is the first to estimate this index in a model of mode choice and commute to different types of job. My index is easy to interpret because I use the linear utility specification of Anderson *et al.* (1992), and introduce value of travel time as a structural parameter.

Figure 1: Proposition 1 - Average trip time larger in case B




Proposition 1 is an implication of the independence of irrelevant alternative (IIA) of the logit model. The IIA property implies that the probability ratio of traveling to any restaurant $i$ and $j$ (see equation 3) depends only on their respective prices, and not on the prices of any other restaurants.

**Proposition 1** *Average trip time ($\bar{t}$) decreases with the addition of a new restaurant at travel time smaller than $\bar{t}$, increases with the addition of a restaurant at travel time larger than $\bar{t}$ and stays constant with the addition of a restaurant at travel time exactly equal to $\bar{t}$.*

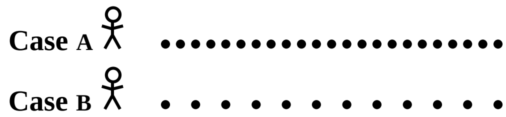
***Proof*** The result follows from the IIA property. Complete proof in Appendix Appendix A. □

The first part of the proposition is a direct consequence of the IIA: the addition of a new restaurant does not change the relative proportions of trips between all other restaurants, so adding a restaurant at exactly $\bar{t}$ cannot change average trip time. The same intuition explains why adding restaurants after $\bar{t}$ increases average trip time, and adding restaurants before $\bar{t}$ decreases average trip time. Figure 1 illustrates an important implication of the proposition. All distances in the figure are in units of time and each dot represents a restaurant in an individual's choice set. In case A, the density of restaurants close to home is high, while the density of restaurants far from home is low, leading to shorter trips, on average, than in case B which displays the reverse density pattern. This result maps into the intuition that someone living in a low-density suburb is likely to travel to the high-density downtown to eat, while a person living downtown is unlikely to drive out to the suburbs for a meal.

The next proposition is the key theoretical result of this section. It formalizes the idea that increasing density is ineffective at reducing trip time, or equivalently that much of the gains from density are gains from variety, that originate from one's ability to visit a preferred location.

**Proposition 2** *Suppose that restaurants are uniformly distributed at distance $t$ from one another. Then as $t$ decreases to 0, $\bar{t}$ converges to a value larger than 0, i.e.* $\lim_{t \to 0+} \bar{t} > 0$.

Figure 2: Proposition 2 - Average trip time similar in case A and case B




***Proof*** In Appendix Appendix A. □

The proposition shows that as density increases to infinity, average trip time does not decrease to zero. In a logit model, individuals make long trips even in the densest areas because additional restaurants never become redundant. In an alternative model with only five different categories of restaurant, $\bar{t}$ would tend to 0 as density increases to infinity and all five categories of restaurants become available infinitely close to home. Simulation results show that $\bar{t}$ converges fast as density increases, and that raising density as in figure 2 has almost no effect on average trip time starting from almost any density level observed in the data.[19] Minimal assumptions about travel demand would imply that increased density fails to induce large reductions in trip time, because individuals choose to gain from density by visiting a location closer to their ideal. The stronger result of Proposition 2 derives entirely from the independence of irrelevant alternatives. The IIA property of the logit and CES models is the object of valid criticism by Ackerberg and Rysman (2005) for instance, but ultimately its relevance is an empirical issue and I postpone this discussion to section 6.

The next proposition shows that moving the entire distribution of restaurants $x$ minutes away from an individual (increasing travel time to each restaurant by $x$ minutes) increases average trip time by approximately $x$ minutes, exactly so if the elasticity of substitution is infinite.

**Proposition 3** *As $\sigma$ becomes large, moving every restaurant $x$ minutes farther from an individual increases $\bar{t}$ by a value arbitrarily close to $x$ minutes as $\sigma$ becomes large.*[20]

***Proof*** In Appendix Appendix A. □

[19]If density is very low, for instance if there is only one restaurant close to home and one very far from home, then average trip time will be almost equal to travel time to the restaurant close to home, because substitution is almost impossible. In this case, increasing density increases travel, because average trip time must increase with the addition of restaurants between the closest and farthest ones.

[20]Simulation results shows that the proposition holds for reasonable values of $\sigma$, in particular for the value that I estimate from the data.

Figure 3: Proposition 3 - Average trip time one minute larger in case A

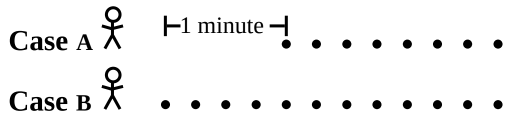



Figure 3 illustrates Proposition 3, which is especially relevant given that a majority of Americans live in residential areas, relatively far from commercial zones. The model predicts that an individual living in a suburb five minutes away from the nearest restaurant drives on average five minutes less on his trip than another individual residing farther outside the city, ten minutes away from the same distribution of restaurants. This result highlights that not all of the gains from density are pure gains from variety, and that residents of dense urban areas may also benefit from shorter trip times, to the extent that they are located closer to a given distribution of restaurants. The next proposition displays some comparative statics.

**Proposition 4** *$\bar{t}$ decreases with the elasticity of substitution $\sigma$, increases with meal price $p$, and decreases with value of travel time $\gamma$.*

***Proof*** In Appendix Appendix A. □

The first part of Proposition 4, that average trip time is low when destinations are highly substitutable, is a core feature of a logit model of travel demand, and it allows to estimate $\sigma$ with travel data. The last two parts of Proposition 4, that average trip time increases with value of travel time and decreases with meal price, correspond to the intuition that long trips are undertaken if travel costs are low or if the object of the trip is valuable. One usually stays close to home to buy a low value item like a coffee, but travels to a preferred and possibly further destination to purchase an expensive watch.

In my presentation of the model, I interpret the elasticity of substitution as a demand-side parameter capturing heterogeneous preferences or a taste for variety. It would also be reasonable to interpret an elasticity of substitution as a supply-side parameter that depends on the diversity of restaurants in an area. In the logit model however, the parameter $\sigma$ is constant across areas and spatial welfare differences come from variation in restaurant density, not in restaurant diversity. In section 7, I propose a simple extension of the model in which restaurants in the same chain are perfectly substitutable. This relaxes the IIA property of the logit model, by taking one aspect of restaurant diversity into account. I also present a nested-logit model, an extension of the logit model in which I introduce a supply-side, and allow restaurants within a given category (e.g. Chinese, Pizza, Sushi) to be more substitutable than restaurants across categories.

## 3. Data

The data on restaurant location come from the Google Places page of each restaurant in the summer of 2011 (these pages are currently called Google+ Local pages).[21] When available, additional data on restaurant characteristics are drawn from Yelp, a popular user review website. The travel data, which identify trips to a restaurant, are from the 2008–2009 edition of the National Household Travel Survey (NHTS). I assume that each traveler resides at his block group's population-weighted centroid, which I obtain from the Missouri Data Center's MABLE Geocorr2K database.[22]

### *Restaurant data*

Data from Google Maps applications offer complete coverage and exact information on restaurant location, both necessary for the innovations of the paper. As an aggregator of local business data, Google Places includes a page for any restaurant with a presence on alternative websites such as Yellow Pages, or an owner willing to create its own page. I collect data on all restaurants in a set of 15 US states containing more than 50% of the US population. I select these states because each of them funded the collection of additional data in my travel database, beyond the federally funded national sample.[23] My restaurant sample consists of 273,000 eating places. Table 1 lists the states in my sample, the number of restaurants I have in each state, and an alternative estimate of the number of 'eating and drinking' places in each state in 2010, from the National Restaurant Association.[24] On average, the two estimates differ by 9%, with a high of +20% in Vermont and a low of -19% in New York.[25] My dataset includes fast food and full-service restaurants, as

[21] I collected the data using PHP, a popular web programming language.

[22] MABLE Geocorr2K computes a block group's centroid from the centroids of each of its constituent census blocks, using census block populations as weights.

[23] I add Arizona, which purchased two regional-level add-on data, but no state-level add-on. The states that I exclude do not have enough travel data to compute estimates of speed at the local level, and too few trips to justify restaurant data collection.

[24] The state-level reports are accessible at http://www.restaurant.org/research/state/. The National Restaurant Association obtains this number from their own research and from federal government data.

[25] I also have a partial sample of 168,000 restaurants in other states of the country (for a total of 440,000 restaurants) to reduce measurement error from trips across state borders.

well as pubs, delis and other eating places. Coffee shops, such as Starbucks, are almost entirely excluded.[26]

For robustness tests and extensions of the model, I also collect data on restaurant characteristics. The Google Places page of a restaurant provides the name of the restaurant and the type of cuisine that it serves (e.g. Korean, American, Chicken, Sushi). I code restaurants into 85 different categories, using definitions from Yelp, the most popular user review website for restaurants. I also identify restaurants belonging to the 50 largest restaurant chains in my sample (the largest of which is Subway). At the time of data collection, about 50% of Google Places pages contained a hyperlink leading to an alternative restaurant page on Yelp. Yelp contains information on average quality ratings from private reviewers (from 0 to 5 in 0.5 increments), prices ($, $$, $$$ and $$$$, that I code as $7.5, $15, $35 and $70),[27] number of reviews, and sometimes on attire (casual 1, dressy 2, formal 3), ambience (9 degrees from casual 1 to upscale 9), parking availability (that I code as as a dummy for the availability of a parking lot, as opposed to only street parking), and whether a reservation is necessary (a dummy variable). A conflict between Google and Yelp occurred about a third of the way through data collection, after which Google removed the link to Yelp from its pages. I therefore have Yelp data for 70,000 restaurants, concentrated in the largest metropolitan areas due both to the data collection strategy and to the geographical preferences of Yelp's contributors.

***Travel data***

The NHTS is a nationally representative survey of travel behavior conducted about every six years. State transportation agencies can also fund the collection of additional (add-on) travel data, which are also publicly available. There are data on 125,000 households in these add-on states I restrict my sample to, representing 90% of the NHTS total. Each participating household member completes a travel diary on a travel day assigned to the household, recording the purpose, length, duration, start time and mode of every trip undertaken that day. Crucially, the data identify trips to 'get/eat meal', the origin of the trips (e.g. home) and the purpose of the next trip (e.g. return home). The data also contain a rich set of individual, household and trip characteristics, as well

[26]Before 2011 Starbucks was not generally considered a restaurant chain, although it is now widely cited as the $3^{rd}$ largest chain in the US. At the time of data collection, Starbucks was not indexed as a restaurant on Google (it is now). Also, trips to coffee shops are identified separately in my travel data.

[27]On Yelp, the dollar signs represent the 'approximate cost per person for a meal including one drink, tax and tip': $ = under $10, $$ = $11-30, $$$ = $31-60, $$$$ = above $61.

Table 1: Number of restaurants in each state

| State | Sample | NRA estimate | Difference (%) |
|---|---|---|---|
| Arizona | 8459 | 9125 | -7 |
| California | 61749 | 62794 | -2 |
| Florida | 32405 | 34914 | -7 |
| Georgia | 15550 | 16081 | -3 |
| Indiana | 10714 | 10875 | -1 |
| Iowa | 5729 | 5908 | -5 |
| New York | 33496 | 41221 | -19 |
| North Carolina | 15763 | 16860 | -7 |
| South Carolina | 8460 | 7652 | 11 |
| South Dakota | 1526 | 1803 | -16 |
| Tennesse | 10136 | 9017 | 12 |
| Texas | 42473 | 37765 | 12 |
| Vermont | 1616 | 1341 | 21 |
| Virginia | 13616 | 13585 | 0 |
| Wisconsin | 9992 | 11961 | -17 |

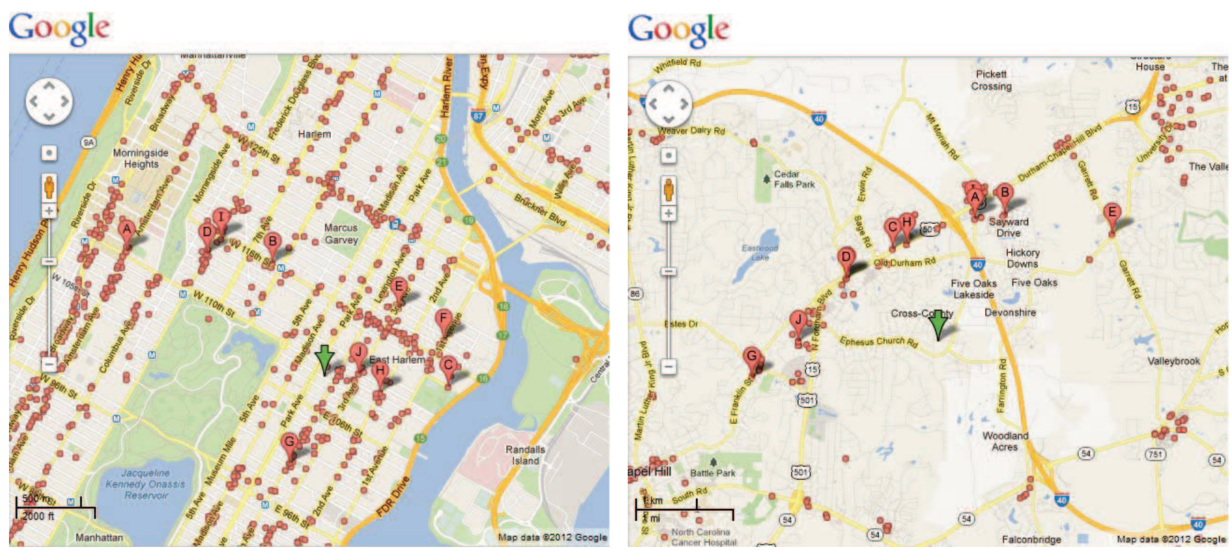
*Notes:* The first column contains the number of restaurants in the sample for each state (data collected in 2011). The second column is a 2010 estimate of the number of 'eating and drinking' places in each state from the National Restaurant Association. The third column contains the percentage difference between the first and second columns.

as the block group in which an individual resides. Trips to or from a restaurant represent 11.5% of all trips, and 26% of households have at least one member going to a restaurant on their travel day. The median trip to a restaurant is about 3 miles and lasts 10 minutes, with higher averages at 6 miles, and 15 minutes. 93% of trips to a restaurants are by privately-operated vehicle ('car', for short) with almost all the remainder by foot. I take walkers into account in the welfare analysis, but for the structural estimation of the model I restrict the sample to trips by car. I also eliminate the small percentage of car trips taken in high-density census tracts in which more than 20% of trips are by foot, because individuals in these areas may choose the car only for long trips, and walk for shorter trips. Limiting the sample in this way ultimately has no effect on any of my estimates.

About 40% of all trips to a restaurants start from home, and for the empirical analysis I restrict the sample to these trips, whose geographical origin is known (at the block group level). Block groups are small, which alleviates concerns about measurement error on the location of a traveler. The median radius of block groups is approximately 0.4 miles, and it is possible to locate individuals where they are more likely to live (the population-weighted centroid). For more than 50% of block groups in my sample, even the restaurant closest to a centroid lies outside of the block group's area.

I remove multiple observations of the same trip for different household members and keep only trips by a driver, leaving 12,000 trips to a restaurant, by a driver, from home. To estimate the logit model, I select 7500 of these trips that are to a restaurant and immediately back home, because trip chaining involves trade-offs beyond the scope of the model. For instance, when a trip to a restaurant is followed by a trip to the movie, the benefits from eating close to the movie theater affect the travel decision, and in that case a long trip does not necessarily imply a willingness to incur travel costs to visit to a preferred restaurant.

*Assembling the data*




(A) A high-density urban area: East Harlem, Manhattan, New York City, NY
(Maps data @2012 Google)

(B) A medium-density suburb in Chapel Hill, NC
(Maps data @2012 Google)

Figure 4: Google Maps with restaurants

*Notes:* Each panel contains a screen-shot from Google Maps resulting from the search command 'Restaurants near [geographical location]'.[28] The downward-pointing arrow indicates the location of an individual's population-weighted block group centroid. Each circle represents the location of a restaurant. The markers from A to G are Google's restaurant recommendations for the search. The scale of the map is at the bottom left. The map in panel B is at twice the scale of that in panel A.

Figure 4 shows two maps, printed from Google Maps. On each map, a downward-pointing arrow indicates an individual's residence, at her block group centroid. Each circle represents the location of a restaurant (the alphabetical markers are Google's recommendations). In the logit model, an individual's restaurant choice set is the set of travel times between an individual and each restaurant in the sample. Travel time, and not distance, is the variable relevant to

[28]Google Maps is available at: `http://maps.google.com/`.

decision-making, because the bulk of all travel costs is value of travel time (Small and Verhoef, 2007). To compute travel time, I first calculate the linear distance between the geographical coordinates of an individual's home and that of a restaurant. I then multiply this linear distance by a correction factor of 1.67, because the driving distance between any two points is longer than the length of the shortest path connecting these points.[29]

Travel time, in minutes, is equal to distance times speed. My measure of speed for each trip depends on the census tract an individual lives in, and on travel distance to the restaurant. To obtain speed estimates, I regress the log of trip speed on the log of trip distance for the entire NHTS sample of car trips, and I include a fixed effect at the census tract level to account for speed differences across areas.[30] Measuring speed as a function of trip distance accounts for the fact longer trips are faster, because they are taken on major urban roads or highways.[31]

### *Variable construction for reduced-form analysis*

I now define variables capturing the features of the restaurant distribution, for use in the reduced-form analysis of the determinants of trip time that I undertake in section 6 to test the logit model of travel demand. To motivate these definitions, it is necessary to highlight three major characteristics of the distribution of restaurants, from the perspective of an individual traveler. To varying degrees, these features are apparent in the two maps from figure 4. The individual in panel A lives in East Harlem, a high-density area in New York City. The individual in panel B lives in a medium-density suburban area of Chapel Hill, North Carolina.

1. *Individuals live relatively far from the closest restaurant.* Most Americans, like the individual in panel B, live in a residential suburb, at some distance from the nearest commercial outlets.

[29] I use a Google Maps application programming interface called Google Distance Matrix to obtain actual driving distance for a representative sample of individual/restaurant pairs (using only the 20 restaurants closest to an individual, which are most relevant). 1.67 is the average difference between the linear distance between two points and the driving distance from Google Distance Matrix. Using the application to compute driving distances (or time) from all individuals to all restaurants in my sample, while feasible, is too costly.

[30] See Couture, Duranton, and Turner (2012) for additional details on these regressions. Note that I drop trips in census tracts for which I have less than 10 trips (6.6% of the sample), as the speed estimates in these areas would be too imprecise.

[31] To facilitate the interpretation of the results and the manipulation of the dataset, I define speed as a characteristic of a geographic area, and I do not let speed vary with individual characteristics. Ignoring variation in speed at the individual level does not affect any of the main results of the paper, but prevents discussion of interesting issues. For instance, Couture *et al.* (2012) find that individuals in lower income households drive slower, which must reduce their welfare gains from variety (see Li (2012) and Handbury (2012) for estimations that take into account the link between income and the gains from variety).

Once the first restaurant is reached, there are many more close to it. Travel time to the closest restaurant is therefore an important descriptive variable.

2. *The number of restaurants available increases more than proportionally with distance (and time)* Both panels suggest that there are many more restaurants available between, say, 10–15 minutes of travel from home than between 5–10 minutes. This result is a geometrical consequence of individuals living on a plane: the area accessible at any given travel distance increases with the square of that distance. That speed increases with travel distance compounds this effect.

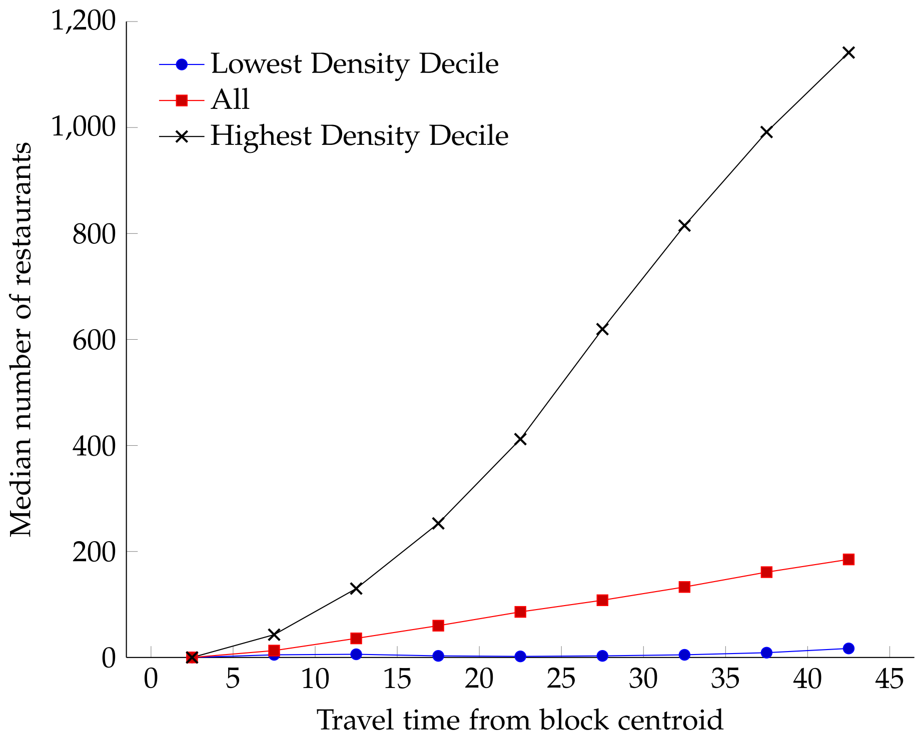
3. *The number of restaurants available increases faster with distance (and time) in dense areas.* In a high-density area (e.g. panel A), restaurants locate on a dense network of major urban roads crossing each other on the plane, and the geometrical argument above fully applies. However, one can think of a very low-density area as a one-dimensional world, in which restaurants locate on the town's sole major road. For instance, one can imagine the individual in panel B driving north to the road on which all restaurants would be located (as shown in the map's corner however, this individual can eventually access many more restaurants farther away). Therefore, within a given travel distance (or time) interval, in dense areas a larger proportion of the mass of restaurants is located far from an individual.[32] This feature of the restaurant distribution is central to the interpretation of reduced-form results on the determinants of trip time.

Figure 5 offers a visual representation of these patterns. I consider each block group in my sample as a different observation. A large sample of block groups is representative of the national diversity of areas, as block groups are census geographic areas covering the entire country and designed to hold similar population (generally between 600 and 3000 people).[33] For illustrative purposes, I only keep a block group if the closest restaurant is within 5 to 10 minutes of travel from its centroid, which is true for 27.5% of blocks in my sample. The middle line in figure 5 plots the median number of restaurants within 5 minute travel time intervals, starting from the centroid and up to 40–45 minutes. For an individual living in these blocks, there are no

[32]If areas with the highest density are also part of the largest continuously populated areas, then this effect is even stronger.

[33] To simplify the dataset, my block group sample consists of 51,641 block groups in which there is at least an NHTS household, representing about 50% of the total number of block groups in the add-on states, and about 25% of the US total.

Figure 5: Median number of restaurants accessible within 5 minute time intervals, computed from the centroid of all block groups for which the closest restaurant is within 5 to 10 minutes of travel




*Notes:* Medians for the full sample are computed from the centroid of all 13881 block groups in which the closest restaurant is within 5–10 minutes of travel (out of 51641 block groups in which there is at least one NHTS household). Medians for lowest/highest density deciles are computed from the centroid of the 1388 block groups with the lowest/highest number of restaurants accessible within 45 minutes of travel.

restaurants within 0 to 5 minutes of travel, so the closest restaurant is relatively far. Past the first location, the number of restaurants increases rapidly with travel time; the median number of restaurants available within 10–15 minutes of travel is 13, and within 40–45 minutes of travel this number increases to 185. The lower and upper lines in the figure show that this increase is even faster in block groups with the highest number of restaurants available within 45 minutes of travel (highest density decile) and considerably smaller for block groups in the lowest density decile.

Consistent with these patterns, I propose four measures of restaurant location that capture the main features of an individual's restaurant choice set.

1. Travel time to the restaurant closest from home.

2. Local density: a measure of local restaurant density passed the closest restaurant. Local density, in restaurants per minutes, is equal to travel time to the $20^{th}$ closest restaurant minus travel time to the closest restaurant, divided by 19.

3. Global density: a measure of restaurant density for an area wide enough to encompass most trips, but not so large as to become irrelevant to a traveler. Global density, in restaurants per minute, is equal to the number of restaurants available within 45 minutes of travel, divided by 45. I also limit the choice set of individuals to 45 minutes in the structural estimation.[34] I experiment with 30-minute and 60-minute choice sets and obtain similar results.[35] 45 minutes corresponds to the $98^{th}$ percentile of trip travel time.

4. Skewness: a measure of whether most of the restaurant mass is distributed close to or far from an individual. This ratio is equal to the density of restaurants from 22.5 to 45 minutes of travel over the density of restaurants from 0 to 22.5 minutes of travel.

Although the simplest version of my model does not allow for variation in restaurant diversity across areas, it is also useful to compute a measure of restaurant diversity. I define diversity as

[34] I could define global density starting starting from the closest restaurant, to be consistent with the theory and the measure of local density. Such a definition would lead to similar regression results and stylized facts, but would make the variable harder to interpret.

[35] There is a theoretical justification for cutting the choice set. In the model of section 2, the addition of restaurants far away from home has almost no impact on average trip time because the probability of visiting them decreases faster than travel time increases. It is easy to show that if restaurants are uniformly distributed and infinitely dense and if $\sigma > 2$, then $\bar{t}$ converges to a finite value as the minutes of restaurants available increases to infinity. This result implies that structural estimation of the model is possible, as otherwise empirical results would depends on how many minutes of restaurants I keep in the choice set.

Table 2: Mean of variables within each decile of global density

| Global density (rest. per minute) | Time to closest rest. (minutes) | Specialization (percent) | Speed (miles/hour) | Pop. density (person/sq. miles) |
|---|---|---|---|---|
| 1.9 | 8.7 | 65 | 11.7 | 70 |
| | (0.9) | (1.3) | (0.3) | (2) |
| 5.1 | 7.1 | 62 | 12.4 | 161 |
| | (0.7) | (0.7) | (0.3) | (4) |
| 8.9 | 6.5 | 60 | 12.4 | 315 |
| | (0.6) | (0.7) | (0.3)) | (7) |
| 13.4 | 5.8 | 58 | 12.3 | 382 |
| | (5.35) | (0.6) | (0.3) | (5) |
| 20.9 | 5.4 | 58 | 12.2 | 517 |
| | (0.4) | (0.8) | (0.3) | (8) |
| 31.5 | 5.0 | 57 | 11.9 | 851 |
| | (0.4) | (0.6) | (0.3) | (17) |
| 45.4 | 4.6 | 55 | 11.7 | 1421 |
| | (0.3) | (0.6) | (0.3) | (23) |
| 65.9 | 4 | 55 | 11.2 | 2190 |
| | (0.2) | (0.6) | (0.3) | (53) |
| 93.8 | 3.7 | 54 | 11.4 | 2772 |
| | (0.2) | (0.6) | (0.2) | (76) |
| 174.9 | 3.1 | 56 | 11.1 | 10045 |
| | (0.2) | (0.5)) | (0.3) | (246) |

*Notes:* Author's computations. Standard deviations in parentheses. Means computed over all 51641 block groups in which there is at least one individual in my NHTS sample. The n$^{th}$ row of the table contains means computed over all block groups in the n$^{th}$ decile of global density. The third column is speed of a one-mile trip. The fourth column is population density at the county level.

a concept separate from that of density. For instance, an urban area containing only Pizza-style restaurants is dense but not diverse, while a sparsely populated area containing many different categories of restaurants is diverse but not dense.

[resume]Specialization: a measure of whether two restaurants in an area are likely to be in the same category. It equals the share of the five most popular restaurant categories among all restaurants available within 45 minutes of travel. The most popular categories need not be the same in each area. This specialization index has the advantage of not being sensitive to the arbitrary definition of smaller categories.

*Summary statistics*

Table 2 contains summary statistics. Each block group in my sample is an observation. The first column shows average global density (the number of restaurants within 45 minutes of travel) within each decile of global density. The other columns contain averages, computed over block groups within each decile of global density, of the following variables: travel time to the closest restaurant, specialization, speed of a one-mile trip and population density at the county level. There is wide variation in global density across areas; it increases by a factor 87 from the first to the last decile. Travel time to the closest restaurant is shorter in denser areas, and it decreases from 8.7 minutes on average within the first decile of global density to 3.1 minutes within the last decile. However, the correlation between travel time to the closest restaurant and global density is far from perfect, reflecting the fact many individuals live in purely residential suburbs at varying driving distance from the same commercial or urban areas. An area's specialization index initially decreases with global density but stabilizes in medium to high-density areas, meaning that restaurants in the lowest density areas are less diverse. Areas in the upper deciles of global density experience slower traffic by about 5–10%, despite the fact higher speed (which is measured with error) increases the distance covered in 45 minutes of travel and therefore has a mechanically positive effect on global density. Finally, there is a strong positive relationship between restaurant density and population density.

## 4. Estimation of the logit model of travel demand

With travel data containing many trips starting from the same location, one can estimate the elasticity of substitution $\sigma$ from a simple OLS regression of the difference in log restaurant prices on the difference in the log share of expenditures on these restaurants. In this case, observing many short trips and few long trips imply that restaurants are very substitutable ($\sigma$ is high) and individuals are unwilling to incur travel costs to visit a preferred destination. In my sample however, almost every trip originates from a different location, and no two areas offer a similar set of available restaurants. I therefore propose a maximum likelihood estimator for $\sigma$ that accounts for the exact restaurant choice set of each traveler.

The estimation sample consists of all trips to a restaurant by a driver that are shorter than 45 minutes, start from home and are followed by a return trip home. Data on trip length is sufficient to estimate the model, as restaurants are only differentiated by their travel time from home, and by a random utility shock. Let $n$ index each trip in my sample of size $N$=7405. Denote the travel time of trip $n$ originating in area $k$ by $m_{nk}$. Recall that the choice set of an individual in area $k$ is $T_k = \{t_{k1}, \dots, t_{ki}, \dots, t_{kI_k}\}$, the vector of travel times to all restaurants within 45 minutes of travel. For about 15% of trips, observed trip time is shorter than my estimate of travel time to the closest restaurant, because of measurement error.[36] In these cases, I assume that a traveler's location is such that $t_1$ is exactly equal to $m_{nk}$, so I add $m_{nk} - t_{k1}$ to the travel time of each restaurant in $T_k$.

The average meal price is constant at $p$ =$13, which is close to the average meal price in the Yelp data. To set a value of travel time, I refer to Small and Verhoeff (2007) who review estimates of the value of driving time from a large literature, and suggest a value equal to 50% of a person's average hourly wage. I set $\gamma$ = 0.2, which corresponds to $12 per hour, or about 50% of the average hourly wage in the United States.[37]

From equation (2), one can write the probability of a trip to a restaurant at travel time $m_{nk}$ as a function of $m_{nk}$, $T_k$, and the parameter $\sigma$:

$$prob(m_n|\sigma, T_k) = \frac{(p + \gamma 2 m_{nk})^{-\sigma}}{\sum_{i=1}^{I_k} (p + \gamma 2 t_{ki})^{-\sigma}}.$$

The maximum likelihood estimate is the value of $\sigma$ that maximizes the probability of observing the sample of trip times $m$ given the choice sets $T$ and the constants $p$ = 13 and $\gamma$ = 0.2.[38] The log-likelihood function is:

$$\ell(\sigma, m, T) = \sum_{n=1}^{N} log(prob(m_{nk}|\sigma, T_k)),$$

and the maximum likelihood estimator is:

$$\hat{\sigma} = \underset{\sigma}{\text{argmax}}\ \ell(\sigma, m, T). \tag{8}$$

I estimate the model by grid-search and find $\hat{\sigma} = 10.5$ (in column 1 of table 3, along with estimates obtained from extensions of the model in section 7). A plot of the log-likelihood function

[36]Such mistakes are often small and occur for short trips to restaurants which are almost closest (e.g. someone enters a 5 minute trip - a round-up value - and I estimate that the closest restaurant is 6 minutes away). There are larger discrepancies in low-density areas with large block groups and imprecise measurement of $t_1$.

[37]My model is consistent with the case of one decision maker, for instance the driver of the vehicle, eating a $13 meal, with a value of time of $12 per hour, and ignoring the preferences of everyone else on the trip when choosing a restaurant. My estimate of $\sigma$ however, turns out to be invariant to the case in which two or more travelers act as one, each with exactly the same preferences, meal price and value of time as the other. That is, I verify that increasing both meal price and value of time by the same factor has almost no effect on my estimate of $\sigma$.

[38]There is never a restaurant located precisely at $m_{nk}$, but this does not affect estimation results. Instead of using the probability of visiting a hypothetical restaurant at $m_{nk}$ for my estimation, I could use the sum of probabilities of visiting all actual restaurants in a small range of minutes around $m_n$ and obtain almost identical results.

Table 3: Maximum likelihood estimation of logit model of travel demand

| | (1) | (2) | (3) | (4) | (5) ) |
|---|---|---|---|---|---|
| $\hat{\sigma}$ | 10.5 | 10.1 | 10.0 | 10.0 | 9.7 |
| | (0.2) | (0.2) | (0.2) | (0.2) | (0.2) |
| $\hat{\beta}$ | | 0.21 | | | |
| | | (0.01) | | | |
| Sorting by value of travel time | | X | | | |
| Meal price varies with distance | | | X | | |
| Perfect substitutability within chain | | | | X | |
| Perfect substitutability within title | | | | | X |
| Observations | 7405 | 7405 | 7405 | 7405 | 7405 |

*Notes:* $\sigma$ is the elasticity of substitution between restaurants, and $\beta$ captures the strength of sorting by value of travel time. Estimates obtained by grid-search in all columns. Standard errors in parentheses computed using the outer-product-of-the-gradient estimator, as suggested in Berndt, Hall, Hall and Hausman (1974).

for different values of $\sigma$ suggests that the function is concave for any reasonable elasticity of substitution, and therefore that $\hat{\sigma}$ is a global maximizer. Estimation results are robust to cutting or expanding the number of minutes in the choice set of individuals, for instance keeping 30 or 60 minutes worth of restaurants leads to $\sigma = 10.3$ and $\sigma = 10.5$. This elasticity of substitution between restaurants is large compared to existing estimates for consumer goods, but it is low enough to generate much extra travel beyond the closest restaurant, and as shown in section 5, substantial welfare gains. I am not aware of other estimates of the elasticity of substitution for services and non-tradables like restaurants.

For a given value of $\sigma$, the model predicts the probability of a trip to each restaurant in each area $k$. With these probabilities, it is easy to compute the predicted time of a trip in area $k$. At $\hat{\sigma}$ = 10.5, the model predicts that the average trip time over all trips in the sample is 13.1 minutes, which is almost equal to 12.9 minutes, the actual average trip time of all corresponding trips in the data (45 minutes and shorter, from home with a return trip). At $\sigma$ = 7.5 or $\sigma$ = 15, the model's prediction of average trip time would be 9.9 and 16.8 respectively, both about 30 standard deviations away from data average. In other words, the maximum likelihood estimator successfully matches the first moment of the trip time distribution. I provide additional evidence on model fit in section 6.

An important concern is that the estimate of $\sigma$ depends on the choice of value of travel time parameter $\gamma$. In their literature review, Small and Verhoef (2007) conclude that depending on individual characteristics, trip purpose and traffic conditions, individual's value of travel time can reasonably range from 20% to 90% of gross hourly wage, with an average around 50%. The 50% average emerges from meta-analyses of hundreds of studies, and as of 2007 was the value recommended by both the US Department of Transportation and Transport Canada. A high end value of $\gamma$ = 0.4, which is close to 100% of the average US hourly wage, leads to $\hat{\sigma}$ = 7.1, and correspondingly higher gains from variety. A very low value of $\gamma$ = 0.1 leads to $\hat{\sigma}$ = 17.4.[39]

An endogeneity problem typically arises when estimating an elasticity of substitution, due to the unobserved relationship between higher prices and better quality. In general, an insensitivity to price differences indicates a preference for higher quality products, not a taste for variety.[40] In my model however, price differences originate from travel costs that are plausibly unrelated to quality differentials. The same restaurant which is located 5 minutes away from an individual is also 25 minutes away for another individual living on the other side of town. Nevertheless, there remain small systematic differences between restaurants that are on average close to travelers and restaurants that are on average far, an issue that I address in the extensions of section 7, using data from Yelp.

## 5. Welfare gains from density

I now turn to estimating the welfare gains from restaurant density. I first compute the variety-adjusted restaurant price index in each area, and convert variation in the index into an average willingness to pay for density. Next, I obtain aggregate measures of the total gains from restaurant density in the US. Finally, I propose a pure travel cost method of evaluating the gains from restaurant density that does not rely on structural estimation of a model.

[39] Individuals who travel to a restaurant have on average higher income, so a value of $\gamma$ = 0.2 may be the on low side. Given my data, it would not be useful to allow $\gamma$ to vary by income, because $\sigma$ is determined by travel costs relative to the price of a meal, and I do not observe the value of the actual meal purchased by travelers, which is likely higher for richer people. The reduced-form results in the next section show that income only has a small effect on travel time, which is consistent with meal price increasing about proportionally with value of time, so that using average values of $\gamma$ and $p$ leads to representative results.

[40] This omitted variable bias is usually dealt with using a technique for panel data pioneered by Feenstra (1994). Identification comes from differences across varieties in the variance of each variety's demand and supply shock, assuming that these shocks are independent over time.

*Welfare differences across areas*

For each block group centroid in my sample, I use equation (6) to compute the variety-adjusted restaurant price index at the maximum likelihood estimate of $\sigma$. The value of the price index at any given location decreases with the number of nearby restaurants, and with travel speed. For a car traveler, the mean and median restaurant price index are both equal to 10, lower than the average price of a restaurant meal ($13) before including transport costs. There is wide variation in the price index across areas. The index ranges from less than 7.5 in much of Manhattan[41] and a few dense parts of San Francisco County with faster car travel, to values above 16 at the $99^{th}$ percentile of the index, in non-metropolitan areas with little gains from variety and hefty transport costs to reach even the closest location.[42] Much of the variation in the index occurs within metropolitan areas. In nine of the ten largest MSAs in my sample, someone moving from a block groups at the $90^{th}$ percentile to a block group at the $10^{th}$ percentile of the price index in that MSA would experience a decrease in the index larger than 20%. The $90^{th}$ to $10^{th}$ percentile differential is lowest in Miami at 14% and highest in Houston and Atlanta at 30 and 32%. These within-MSA variations in the index reflect the very localized nature of the gains from restaurant density. Remote restaurants, that are expensive because of travel costs, have little impact on welfare because the elasticity of substitution between restaurants is high. For instance, preventing access to restaurants between 30 and 45 minutes of travel would reduce the average price index by only about 2%.[43] The conclusion that the gains from restaurant density are localized may generalize to much of the consumption benefits of density, given relatively short trip times for many other kinds of non-work trips in the NHTS (e.g. medical/dental, grooming/haircut/nails, attorney/accountant, gym/exercise/play sports, etc).

A decline in the variety-adjusted restaurant price index can translate into sizable welfare gains for an average household. I first compute the monetary value of the gains from density in partial equilibrium, and then from an extension of the model allowing for substitution between restaurants and all other goods. To convert the relative price index of equation (5) into a willingness to

[41]Not taking parking costs into account leads to significant underestimation of the index for a car traveler in Manhattan and other very high density areas. However, as I show in section 8, a walker in Manhattan faces an index close to 7.5, and still lower than that almost anywhere else in the US.

[42]My sample covers only 50% of the US population, but it is unlikely that areas in Chicago or Philadelphia which are large and dense cities excluded from my sample, have a lower price index than Manhattan.

[43]Removing restaurants far away has a larger effect on trip time; average trip time is 14.83 minutes over all trips from home by a driver, and it drops to 13.47 keeping only trips shorter than 45 minutes, and to 12.53 minutes keeping only trips shorter than 30 minutes.

pay for density, consider a move from an area $k$ to a counterfactual area $k'$. From equation (4), the expected value of the indirect utility in area $k$ is $v_k = ln(y_k P_k^{-1})$, where $y_k$ is expenditure on restaurants in area $k$. Define $y_{k'}$ as the value of restaurant expenditures that makes the mover indifferent between living in area $k$ or $k'$, so that $y_{k'}$ is such that $ln(y_k P_k^{-1}) = ln(y_{k'} P_{k'}^{-1})$. Some easy algebra leads to:

$$y_k - y_{k'} = y_k(1 - P_{k,k'}). \tag{9}$$

I take expenditure shares from the Consumer Expenditure Survey (CEX) 2009, in which food away from home represents on average 5.3% of household expenditure, or $2600.[44] Using these numbers, an average household's willingness to pay to prevent a 20% increase in the restaurant price index is |2600*(1-1.2)|=$520 annually. In section 6, I use regression analysis to show that most of the welfare gains from an increase in restaurant density are gains from variety, as opposed to gains from travel time savings through shorter trips.

This simple computation may overestimate the gains from density for three reasons. First, it does not allow for substitution between expenditures on restaurants and expenditures on all other goods. In Appendix Appendix C, I find that a 20% increase in the price index leads to a 17% decrease in the probability that an individual travels from home to a restaurant on any given day. This decrease does not suggest a very large price elasticity of demand for restaurants, and most of it is compensated by an increase in restaurant trips originating from somewhere other than home. To measure exactly how substitution of restaurants for all other goods affect my estimates, I specify a nested-logit model with two nests, one for restaurants and one for all other goods (the details are in Appendix Appendix D). I set the price elasticity of demand for restaurants at -1, consistent with a literature review by Okrent and Alston (2010) who find an average value of -1.02 for the price elasticity of demand for food away from home. Using this model, the average household's willingness to pay to prevent a 20% increase in a restaurant price index drops to about $475.[45] Second, I may overestimate the gains from density because unless every passenger on a restaurant trip has the same preference, the gains from restaurant density are lower for a passenger who is not part of the travel decision. Assuming that only trip drivers decide what

[44] Unlike my index, the CEX considers transport costs separately from other expenditures. Also, expenditures on food away from home in the CEX includes coffee, ice cream and snacks, which my index does not take into account.

[45] Okrent and Alston (2011) find an very high price elasticity of demand for full-service restaurant (close to -3), but with much substitution to limited service restaurants, which account for a slightly larger share of restaurant expenditures. Using a high price elasticity of demand for restaurant of -2 instead of -1, I find that for an average household's gains from density are lower by another 18%. This suggests that my estimates are of the same order of magnitude for reasonable value of the price elasticity of demand for restaurants.

restaurant to visit and benefit from density reduces total gains from density within households by 37%.[46] Third, I compute the index for trips starting exactly from home, which is true of 40% of all restaurant trips. Other restaurant trips may originate from higher density areas, especially the 10% of trips starting from work.

It is interesting to compare these results with those of Handbury and Weinstein (2012), who estimate a variety-adjusted price index for tradable consumer goods (groceries) in different metropolitan areas. They find that residents of larger cities, controlling for store amenities, individual characteristics and differences in the number of varieties available, face a lower price index for groceries. This price index drops by 10% from New York City to Des Moines, the smallest city in their sample. They estimate that New York residents have access to 97,000 types of groceries within the MSA, versus 32,000 in Des Moines, leading to a 3% decrease in the price index that is purely due to variety. While these numbers are estimated for large areas and do not take transport costs into account, simple comparisons with my results for the restaurant industry suggest much larger spatial welfare differentials in the non-tradable service sector. Many residents of the New York or Los Angeles metropolitan areas have access to more than 10,000 restaurants within 45 minutes of car travel, and would face at least a 25% reduction in the price index from moving anywhere in Des Moines, a city with faster car travel but with only 650 restaurants available within 45 minutes of travel.

Future research may demonstrate that individuals derive similar benefits from the higher density of health providers, entertainment options and other services in the downtown cores of metropolitan areas. Part of the explanation for the importance of non-tradables in accounting for the consumption advantage of dense areas is the highly developed supply chains of the major consumer-goods retailers. Thanks to the low cost of moving goods, residents of America's suburbs and smaller cities have access to an impressive variety of tradables. Such feat is not easily replicated in the non-tradable service sector, which depends to a larger extent on the movement of people. Dense urban areas still have a unique advantage in reducing transport costs between individuals.

[46] 63% of individuals taking trips to restaurants from home by car are identified as drivers in the NHTS. Note that this result is inconsistent with average of the NHTS variable for the number of individuals on each of these trips, which equals 2.2.

*Aggregate gains from density*

I now propose three counterfactual experiments to provide measures of the aggregate gains from restaurant density in the United States. In the first experiment, I compare the area $k$ an individual lives in, with a counterfactual area $k'$ in which only the closest restaurant is available. This experiment attempts to isolate the part of the gains from density that is purely due to gains from variety. It holds travel time to the closest restaurant constant, and only measures the benefits from access to preferred destinations; from not having to always visit the same location. The first row of table 4 contains the average value of the relative price index for this experiment, over all trips from home by a driver. The average value of the index is about 1.5 in the first column for the logit model, with slightly higher values in other columns containing results for the model's extensions of section 7. The index ranges from 1 to almost 2 in very high density areas. In locations with the largest gains from variety, restaurant prices would have to double to make an individual indifferent between sustaining this price increase and being forced to always visit the closest restaurant (holding expenditures on restaurants constant). While there is much empirical evidence to support setting the value of travel time at $\gamma = 0.2$, the welfare estimates are sensitive to this choice. In the logit model, using $\gamma = 0.1$ leads to an average relative price index of 1.27 and using $\gamma = 0.4$ leads to 1.88.

As before, I compute the monetary value of these gains from variety first in partial equilibrium, and then from an extension of the model allowing for substitution between restaurants and all other goods. For a traveler, the daily expenditures on restaurant $y_k$ are equal to meal price plus the value of transport costs for the trip, and the welfare gains from variety is given by the absolute value of $y_k(1 - P_{k,k'})$, where $P_{k,k'}$ is the relative price index. I use the NHTS sampling weights to sum these individual gains over all trips to a restaurant by a driver (including trips not from home and by foot) and then multiply by 365 to obtain an annual value.[47]

[47] To obtain numbers valid at the national level, I exploit features of the travel survey's design. The NHTS is representative of travel behavior in the entire country. Each participating individual is assigned a weight, equal to the number of Americans that he or she represents (so the weights sum up to the US population). My sample represents around half of the US population, so I scale up these weights by a factor of about two. Using these weights, I can compute aggregate values for travel time, expenditures on restaurants or gains from variety. In these aggregate computations, I include all trips to a restaurant by car or foot, including trips not from home, and excluding only the negligible fraction of trips by public transportation Whenever an individual travels by foot, I set a constant travel speed at 15 minutes per mile to measure travel time to restaurants in her choice set (the NHTS suggests 12.5 minutes, and Google Maps usually assumes walking speed of around 20 minutes per mile). I assume that all trips start from home, although the restaurant choice set at the true origin of some trips is not exactly the same as that of a trip starting from home. In line with regression results in the next section, I apply a -20% correction factor to trip time for all trips that are part of a chain, because restaurant preferences only partially explain longer trips in this case.

Table 4: Average restaurant relative price indices

| Average relative price index | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Counterfactual is closest restaurant only | 1.495 | 1.549 | 1.531 | 1.504 | 1.527 |
| | (0.126) | (0.137) | (0.141) | (0.098) | (0.133) |
| Counterfactual is $5^{th}$ percentile price index | 0.850 | 0.841 | 0.846 | 0.841 | 0.843 |
| | (0.099) | (0.102) | (0.103) | (0.098) | (0.101) |
| Counterfactual is $95^{th}$ percentile price index | 1.400 | 1.414 | 1.413 | 1.415 | 1.410 |
| | (0.152) | (0.161) | (0.173) | (0.181) | (0.159) |
| Sorting by value of travel time | | X | | | |
| Meal price varies with distance | | | X | | |
| Perfect substitutability within chain | | | | X | |
| Perfect substitutability within title | | | | | X |
| Observations | 13830 | 13830 | 13830 | 13830 | 13830 |

*Notes:* Computed using all trips from home by a driver (13830 trips (including all trips in high-density tracts and in areas with imprecise speed estimates). Standard deviations of the relative price index (not of its mean) in parentheses. NHTS sampling weights used to compute all means and standard deviations. Because of computing limitation, I use the counterfactual $5^{th}$ and $95^{th}$ percentile areas of the logit model without sorting (column 1) to compute relative price indices in the model with sorting (column 2).

I find aggregate gains from variety of $160 billion, or about 2.7% of consumer expenditures. Assuming that only drivers gain from variety reduces these gains by 37%[48]. Accounting for substitution between restaurants and all other goods leads to a further drop of 18%, in part because this computation requires use of CEX data. That is, using the NHTS travel data I compute that individuals in 2008–09 spent $348 billion per year eating out, while according to the CEX households spent a lower 323 billion on food away from home in 2009. I conclude that there are large gains from access to restaurants beyond the closet eating option available. The most conservative estimate of these gains from variety is at $80 billion (25% of restaurant expenditures and 1.4% of consumer expenditures), with an upper bound at $160 billion (50% of restaurant expenditure and 2.7% of consumer expenditures).[49]

In the two other experiments, the counterfactuals are existing areas. I rank each block group in my sample by the restaurant price index in its centroid. In the first experiment, I compare the block group an individual lives in to a denser counterfactual block group at the $5^{th}$ percentile

[48]I assume that as with car travelers, 63% of walkers are decision makers.

[49]I compute total consumer expenditures from the CEX as average annual household expenditure ($49,056) times total number of households (120,847,000). Household expenditures do not include the value of travel time, and adding it would make a small difference in the percentage that I estimate. For instance if I assume that each household drives 2 hours per day, valued at $12 per hour, and add this travel costs to household expenditures, $100 billion drops from representing 1.68% of total expenditures to representing only 1.43%.

price index. The average value of the relative price index is equal to 0.85 (row 2 of table 4). In the second experiment, I compare the area an individual lives in to a lower density counterfactual area at the $95^{th}$ percentile price index. In this case the average relative price indices is between 1.40 (row 3 of table 4). I conclude that the gains to individuals from moving into the densest areas are small relative to their losses from moving into the lowest density areas. Note that only a part of the welfare gains from moving out of very low density areas are gains from variety. In the countryside, even the closest restaurants are located far from travelers, which necessarily generates welfare losses from longer trip times.[50]

***A pure travel costs method***

A simple method of assessing the magnitude of the gains from variety in the restaurant industry is to compute the cost of all travel passed the nearest destination that would satisfy the objective of a trip. If, on his way to his preferred restaurant, an individual incurs $2 in transport costs beyond the closest available location, then the value of access to a variety of destinations must be at least $2. This computation, which does not require the structural estimation of a model, provides a lower bound for the (gross) aggregate gains from variety.

I first compute the difference between actual trip time and travel time to the closest restaurant for all trips to a restaurants.[51] Summing up this difference over all trips using NHTS weights and multiplying by 365 days per year, I find that Americans spend about 5.6 billion hours traveling beyond the restaurant closest to their home. At $12 per hour, this travel time is worth nearly $67 billion dollars; a direct estimate of the minimum amount that Americans are willing to pay every year to eat at a restaurant that they prefer. This implies that individuals acquire the net gains from variety, that I estimate at $80–160 billion, through an investment in travel time of $67 billion. This sizable return on travel (more than 100%) must be given due consideration when designing tax policies to reduce private travel.

It is also instructive to consider non-structural evidence that the gains from variety are greater in denser areas. I rank each trip in my sample by global density in the location it originates from. The median number of restaurants passed before reaching destination is 14 within the first

[50] I provide more precise results on the relationship between density and trip time in the next section, but note that the median length of a trip to a restaurant is 10 minutes and that most individuals living in block groups at the $5^{th}$ percentile of the index do not have access to any restaurants within 10 minutes of travel.

[51] As before, I assume that an individual travels to the closest restaurant whenever time to the closest restaurant is larger than trip time due to measurement error.

decile of global density, and rises steadily up to 89 in the last, denser decile (means are larger than medians). These numbers strongly suggest that travelers in dense areas visit, on average, destinations closer to their ideal. These computations also hint at the reduced-form regression results of the next section on the determinants of trip time: if the number of restaurants passed on each trip increases with density, then raising density will not lead to a large reduction in trip time.

## 6. Testing the model: Reduced-form regressions on the determinants of trip time

In this section I test the logit model of travel demand. Of particular interest is the prediction that if individuals benefit from density by visiting preferred destinations, as the model's structural estimates of the gains from density implicitly assume, then increasing restaurant density should have little impact on travel time. First, I show how to compute the predicted value of average trip time for each traveler in my sample. I then run reduced-form regressions of trip time on measures of restaurant location using both predicted and actual trip time as a dependent variable. If these two sets of regressions generate similar coefficients, then the model accurately predicts the effect of measures of restaurant density on trip time.[52] I also propose an instrumental variable strategy, to investigate omitted variable issues. In Appendix Appendix C, I present complementary results on the determinants of the probability of making a restaurant trip.[53] Of course, the regressions on the determinant of trip time in this section are also interesting in their own rights, for instance the relationship between travel and the built environment is crucial when evaluating urban renewal schemes designed to reduce vehicle miles traveled.

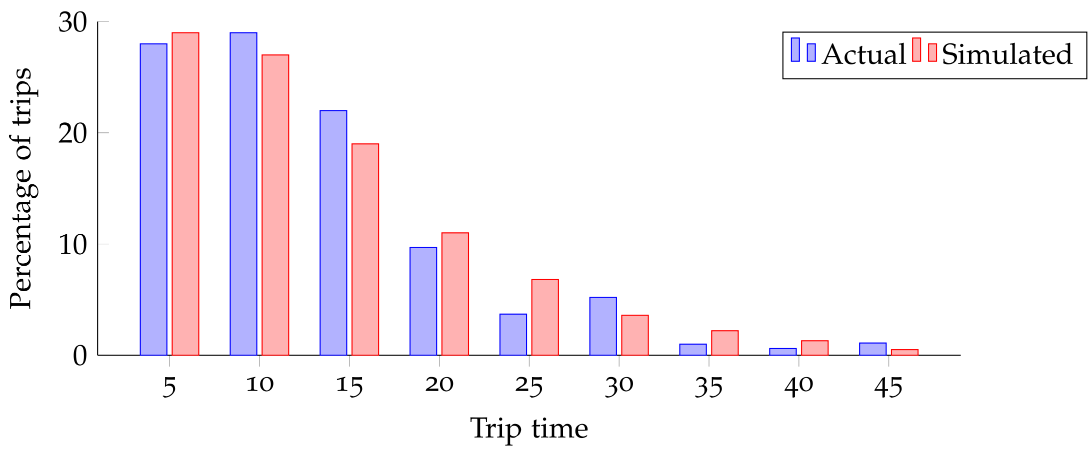
### *A simulated trip dataset*

Before the regression analysis, it is instructive to visually compare the distribution of trip times in the data with a distribution of simulated trip times generated by the model. The model predicts the probability of traveling to any given restaurant from each location, given the set of travel

[52]Note that variation in restaurant density across areas is not what identifies the parameter of the model (see Proposition 2), so that the model does not necessarily match the effect of density on trip time across areas. Therefore, the regression coefficients that I report do not depend on the particular value of the elasticity of substitution that I use to obtain predicted value (except for the constant term).

[53]Other studies find that the socio-economic characteristics of individuals, as opposed to the built environment, are the main determinants of the number and types of trips that they take. I confirm this finding and find that income is by far the most important predictor of whether an individual eats out, with rich people having a considerably higher probability of traveling to a restaurant.

Figure 6: Distribution of trip times, actual and simulated data




*Notes:* Sample of 7303 trips less than 45 minutes, by a driver, from home and followed by a return trip. The y-axis represents a time interval around the values indicated. This interval is 0 to 7.5 minutes for 5 minutes trip, 42.5 to 45 minutes for 45 minutes trip, and otherwise $x - 2.5$ to $x + 2.5$ for a $x$ minutes trip. Simulated data obtained using the maximum likelihood estimate of $\hat{\sigma}$ = 10.5.

times to all restaurants and an estimate of the elasticity of substitution. To obtain a simulated trip dataset, I draw a trip time for each driver in the sample (round-trips from home, shorter than 45 minutes), using the probability distribution given by the model. For a given estimate of $\sigma$, this simulated dataset is a sample of trip times that the travel survey could have collected if the logit model had generated the data.

Figure 6 shows the proportion of trips within 5 minute time intervals from 0–5 to 40–45 minutes, in both simulated and actual data. The logit model does remarkably well in matching the distribution of trip times, considering that only one parameter ($\sigma$) is estimated to fit the data. The proportion of 5-minute trips is almost the same as that of more expensive 10-minute trips because many travelers do not have any restaurants available within 5 minutes of travel from home.

***OLS regressions using trip time predicted by the logit model***

Knowing the probability of a trip of each length in an area $k$, I can compute $\bar{t}_{nk}$, the model's prediction of expected trip time for each trip $n$ in the sample. I now run OLS regressions using $\bar{t}_{nk}$ as a dependent variable. These regressions on the determinants of average predicted trip time reveal the impact that measures of restaurant location would have on trip time if the model were true (the propositions of section 2 are informative, but only valid for hypothetical restaurants

Table 5: The determinants of trip times, predictions from the logit model

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| log Predicted average trip time ($\bar{t}$) | | | | | |
| log Global density | $0.103^a$ | | $0.144^a$ | $0.135^a$ | $0.220^a$ |
| | (0.004) | | (0.002) | (0.003) | (0.003) |
| log Skewness | | $0.113^a$ | | $0.022^a$ | $-0.055^a$ |
| | | (0.004) | | (0.003) | (0.003) |
| log Time to closest rest. | | | $0.485^a$ | $0.468^a$ | $0.540^a$ |
| | | | (0.006) | (0.007) | (0.006) |
| log Local density | | | | | $-0.230^a$ |
| | | | | | (0.003) |
| Observations | 7406 | 7397 | 7406 | 7397 | 7397 |
| $R^2$ | 0.12 | 0.15 | 0.65 | 0.65 | 0.80 |

*Notes:* OLS regressions with a constant in all columns. Robust standard errors in parentheses. $a$, $b$, $c$: significant at 1%, 5%, 10%.

distributions). The estimating equation is:

$$log(\bar{t}_{nk}) = \alpha + \beta_1 loc_k + \epsilon_k, \tag{10}$$

where $loc_k$ represents measures of restaurant density in area $k$ (travel time to closest restaurant, global density, local density, and skewness). I estimate this equation on a sample of 7406 observations, indexed by $n$, and corresponding to each actual trip to a restaurant by a driver, from home and back and shorter than 45 minutes. Table 5 contains the regression results.

All the coefficients in table 5 are elasticities.[54] The four measures of restaurant location are correlated with one another, so the coefficient on each variable is sensitive to the inclusion of the others. The predicted effect of global density on trip time is positive, with an elasticity ranging from 0.10 in column 1 when the variable enters alone, to 0.22 in column 5 when all variables are included. This corresponds to trips about 2 to 4 minutes longer than average in the densest decile of global density. According to Proposition 2 and 3, the effect of global density on trip time, controlling for time to the closest restaurant, would be about zero if restaurants were uniformly distributed. So Proposition 1 implies that this positive coefficient is due to the larger share of the restaurant mass located far from home in dense areas (see section 3 for a precise discussion). In column 2 the skewness of the restaurant distribution enters the regression alone, and its coefficient is positive at 0.11, so the model predicts that trips are longer in areas with a

[54] I also run regressions (results not shown) using dummy variables for the deciles of global and local density, and in which travel time to the closest restaurant and my measure of skewness of the restaurant distribution enter linearly.

disproportionately large number of restaurant far from home. This ratio of density close and far from home, however, is a coarse measure of location and its effect is reduced or negative when global and local density enter the regression. Unlike the effect of global density, the effect local density is negative (elasticity of -0.23, in column 5). High density close to home implies relatively lower density far from home, hence the shorter predicted trip times. The effect of time to the closest restaurant is harder to interpret as an elasticity, but its linear effect is close to 1 at about 1.2.[55] So the model predicts that an individual living 15 minutes away from the closest restaurant on average makes a trip about 5 minutes shorter than that of another individual who lives 20 minutes away from the closest restaurant. These four measures of restaurant location capture most but not all of the features of the distribution of restaurants that are relevant to the model. The $R^2$ of the regression with global density as the only regressor is 0.12, adding a measure of skewness does not make much difference, adding time to the closest restaurant increases $R^2$ to 0.65 and adding local density increases $R^2$ to 0.80.[56]

***OLS regressions using actual trip time***

For regressions on the determinants of actual trip time, I keep the sample of 11900 trips to a restaurant from home, by a driver. I verify that keeping only the 7406 trips followed by a return home and shorter than 45 minutes (to match the regressions on predicted average trip time) leads to similar results. The estimating equation is:

$$log(m_{nk}) = \alpha + \beta_1 loc_k + \beta_2 div_k + \beta_3 X_n + \beta_4 Z_n + \epsilon_{nk}. \tag{11}$$

Each observation $n$ is a trip. The dependent variable $m_{nk}$ is travel time of a trip in minutes. $loc_k$ represents measures of restaurant location in the area $k$ from which trip $n$ originates. $div_k$ is a control for restaurant diversity, for instance the specialization index. $X_n$ is a vector of individual characteristics for the driver of trip $n$, such as age, gender, education, and income (I omit to index individuals, who rarely take more than one trip to a restaurant on their travel day). $Z_n$ is a vector of trip characteristics, including for instance the number of individuals on the trip, the start time,

[55] Proposition 3 predicts an coefficient equal to 1. The effect of time to the closest restaurant in the regression of table 5 is probably overestimated, as it is the only variable that measure the exact location of a restaurant (other variables are proxies), and it captures other features of the distribution.

[56] Regressions on a dataset of simulated trips (instead of predicted average trip time) leads to similar coefficients but much lower $R^2$ (which would drop from 0.8 to 0.2 in the regression of column 4). The lower $R^2$ in regressions on simulated trip time reflects the randomness inherent in the logit model of travel demand; the model cannot accurately predict the length of any single trips, as destinations close and far are equally likely to receive a large idiosyncratic shock.

Table 6: The determinants of trip time

| Trip time to restaurant | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| log Global density | -0.015$^b$ | 0.010$^c$ | -0.015$^b$ | 0.033$^a$ | 0.040$^a$ | 0.033$^a$ | 0.030$^a$ | 0.007 | -0.002 |
| | (0.006) | (0.006) | (0.006) | (0.008) | (0.009) | (0.008) | (0.008) | (0.009) | (0.017) |
| log Skewness | | | 0.060$^a$ | 0.016$^b$ | -0.01 | 0.014$^c$ | 0.018$^b$ | 0.015$^b$ | 0.007 |
| | | | (0.006) | (0.007) | (0.009) | (0.007) | (0.007) | (0.007) | (0.009) |
| log Time to closest rest. | | 0.237$^a$ | 0.188$^a$ | 0.231$^a$ | 0.244$^a$ | 0.231$^a$ | 0.220$^a$ | 0.210$^a$ | 0.203$^a$ |
| | | (0.013) | (0.013) | (0.014) | (0.017) | (0.014) | (0.014) | (0.014) | (0.017) |
| log Local density | | | | -0.130$^a$ | -0.112$^a$ | -0.130$^a$ | -0.129$^a$ | -0.132$^a$ | -0.145$^a$ |
| | | | | (0.011) | (0.013) | (0.012) | (0.011) | (0.011) | (0.013) |
| log Specialization | | | | | | | | -0.290$^a$ | -0.079 |
| | | | | | | | | (0.087) | (0.25) |
| **Controls** | | | | | | | | | |
| Individual characteristics | | | | | | X | X | X | X |
| Trip characteristics | | | | | | | X | X | X |
| MSA fixed-effects | | | | | | | | | X |
| **Sample** | | | | | | | | | |
| Round-trips only | | | | | X | | | | |
| Observations | 11865 | 11865 | 11828 | 11814 | 7397 | 10846 | 10659 | 10331 | 8759 |
| $R^2$ | 0.0006 | 0.032 | 0.040 | 0.051 | 0.041 | 0.063 | 0.157 | 0.158 | 0.157 |

*Notes:* Regressions with a constant in all columns. Robust standard errors in parentheses. $a$, $b$, $c$: significant at 1%, 5%, 10%. Dependent variable is trip time to restaurant in all columns. Individual characteristics include 17 dummies for household income, 4 dummies for education, household size, 6 dummies for age, a dummy for gender, a dummy if black and a dummy for worker's status. Trip characteristics include 5 dummies for each peak hour (7-8am, 8-9am, 15-16pm, 16-17pm, 17-18pm), a dummy for trips followed by a return trip home, a dummy for trips on week-end, the number of children on the trip, the number of adults on the trip, the number of non-household members on the trip, the log time spent at destination, the log speed of a one mile trip in the driver's census tract, and the the log of gas price in the driver's region on the week the trip was taken. The sample size is lower than that in table 4, because here I eliminate trips in census tracts with more than 20% walking trips, or in which travel speed has to be estimated at the county level. In the 'Round-trips only' row, I limit the sample to trips to a restaurant which are immediately followed by a return trip home.

the time spent at destination, a dummy for whether the trip is followed by a return trip back home, and travel speed in area $k$. The regression results are presented in table 6.

Columns 1–5 present regression results with only measures of restaurant location as controls. Column 5 excludes all trips not followed by a return home – which barely affects the coefficients – and so is directly comparable to column 5 of table 5. The effect of global density on trip time is close to zero in each regression, meaning that the number of restaurants available within 45 minutes of travel has little impact on trip time. That increasing global density fails to reduce trip time in the data corroborate the intuition behind the model, that substituting among travel destinations is cheap in dense areas, so individuals often gain from density by visiting preferred location. However, this zero effect does not match the model's prediction of a positive effect, and this discrepancy casts doubts on the logit model's strong assumptions. If the IIA does not hold for instance, then additional restaurants eventually become redundant and the mass of restaurants far from home in dense area may exert little attraction on a traveler. The model's extensions address these issues.

The elasticity of trip time with respect to travel time to the closest restaurant is large and positive, ranging from 0.2 to 0.25 depending on the regression, but it is smaller than predicted. If time to the closest restaurant entered linearly, it would have an effect of about 0.5.[57]

The elasticity of trip time with respect to skewness is 0.06 in column 3, in a regression also including time to the closest restaurant and global density as controls. That is, individuals take shorter trips if the mass of restaurant within 45 minutes is disproportionately located close to home. Most of the effect of skewness disappears when controlling for local density in column 4. Local density has the predicted negative effect on trip time (elasticity of -0.13), so travelers with a high density of restaurant close to home (passed the closest restaurant) make shorter trips. Proposition 1 can explain the estimation results in column 4, in which global density a small but positive effect and local density has a negative effect: for individuals living in areas with the highest density of restaurants close to home, the density of restaurant far from home is necessarily lower in relative terms.[58] A general conclusion is that while increasing the level of density over a large area has little effect of trip time, the decision of how far to travel strongly depends on the distribution of restaurants in space.

[57]Some of this discrepancy is due to measurement error. The linear effect of distance to the first restaurant on trip distance is larger, between 0.6 and 0.9, as distance does not suffer from measurement error in speed and trip distance seems to be entered more precisely in the travel data. The distance to the first restaurant is itself measured with error, because the location of each individual is only known at the block group level. If this error is more likely to be negative, because individuals in large unpopulated blocks live closer to commercial centers than my measure of block centroid suggests, then my estimate is biased downward. By replacing time to the closest restaurant with trip time whenever the latter is smaller, as I do to estimate the model, I obtain linear coefficients of about 1. I conclude that travel time to the first restaurant has a large and positive effect on trip time, and that even more precise data may confirm the theoretical prediction of a proportional effect.

[58]The variables for local density and time to the closest restaurant are a measure of what Ewing and Cervero (2010) call 'destination accessibility' in their meta-analysis. It is often measured as distance to downtown in studies on the determinants of vehicle miles traveled, and it has a robust negative effect on travel. My model suggests that this finding should be understood in relative, not absolute terms. That is, individuals living downtown make shorter trips in part because there are much less options farther out, and increasing suburban restaurant density could have a positive effect on trip length in the city center. Kockelman (2001) also shows that individuals substitute destinations close for destinations far, by estimating a demand system in which individuals can choose between different travel zones.

In column 6, I add individual characteristics to the regression. These additional controls do not affect the coefficients on measures of restaurant location, and have little explanatory power. The lowest income households (less than $20,000 per year) take trips about 10% shorter than the highest income households (more than $100,000 per year), but this effect is no longer significant when controlling for trip characteristics. There is no difference in trip time for families with income between $20,000 and $100,000. Education has very little impact. Trip time increases with age, as driver's in their 20s and 30s take trip about 20% shorter than those of drivers 75 and older. Members of larger households take slightly shorter trips, and this effect is entirely due to children, as additional adult members have a small positive effect. Blacks take trips 10 to 15% longer.[59] Female drivers take trips about 2% shorter, but the result is at the margin of significance. Worker status has no significant effect on trip time. Given that the maximum likelihood estimator of the elasticity of substitution essentially matches the mean of the trip time distribution, these OLS regression results suggest small differences in $\sigma$ across groups of individuals with different characteristics. In this paper however, I only estimate $\sigma$ for the whole sample, and focus on geographical differences in the welfare gains from variety.

Controls for trip characteristics, included in column 7, have more explanatory power but also do not affect coefficients on measures of restaurant location. Trips not followed by a return home are 20% longer, consistent with the idea that individuals undertake a trip chain based on their preferences for two or more objectives, not just eating. Doubling the time spent at destination increases travel time by 12%. Trips during the morning peak are slightly shorter, perhaps because breakfast is a standardized and inexpensive meal. Trips during the evening peak - worse in terms of traffic - are almost 15% longer. The dummy for week-end trips is not significant. Traveling with children has no impact on trip time, but each additional adult on a trip increases its length by 9%. Whether or not the passengers are household members has no effect on trip time. This relatively small but positive impact of each passengers on trip time may be due to sharing of gasoline costs, but it can also denote joint decision making by individuals with different preferences maximizing an objective function. A discrete-choice model relaxing the assumption of a single decision maker,

[59]In a regression of trip distance on measures of density in restaurant per miles and individual characteristics, a black dummy has almost no effect. The reason is, shorter trips for blacks are mostly due to slower driving speed. Similarly, more than half of the age effect disappears in regression with trip distance as the dependent variable, and the poorest households make trips that are 30% shorter in units of distance (instead of 10% shorter in units of time), consistent with Couture *et al.* (2012) who find that driving speed increases with income. Recall that the measures of density in restaurant per minutes do not account for variation in driving speed at the individual level.

which implies that passengers should not affect trip time, may be a better fit for the data, and could be an avenue for future research. The elasticity of gasoline price on trip time is about 0.03, at the margin of significance.[60] Increasing travel speed in the area in which a trip originates slightly reduces trip time.[61]

In column 8, I add the specialization index, a measure of restaurant diversity, as a control. High specialization of restaurant categories in an area has a small negative effect on trip time. In a regression of trip time on 9 dummies for the deciles of specialization and other measures of restaurant location, I find that trip time is about 1 minute shorter in the 5 deciles with highest specialization, a result not robust to the inclusion of individual and trip characteristics. Trip time does not change from low to medium levels of restaurant specialization. Including specialization decreases the effect of global density from 0.03 to 0, so the effect of global density on trip time would be smaller if restaurants were not more diverse in dense areas. The correlation between specialization and global density (and other measures of location) mean that this effect is hard to estimate precisely, and can be due in part to multicollinearity. In the logit model, restaurant diversity is constant across areas, but the idea that individuals would travel less if destinations were identical is at the core of any model with substitutable destinations. The nested-logit model of section 7 can explain this small but positive effect of restaurant diversity on trip time.

In column 9, I add Metropolitan Statistical Areas (MSA fixed-effects and exclude trips in non-MSA areas. Results are similar, except for the effect of specialization which becomes insignificant, probably because most restaurants within 45 minutes of travel from home are also within MSA boundaries, and because specialization only has a sizable impact in low diversity, generally out-of-MSA areas.

To summarize, the logit model describes the first-order features of the data, but there is a sizable discrepancy between the actual and predicted effect of the number of restaurants within 45 minutes of travel (global density) on trip time. I offer three explanations. First, remote restaurants may be close substitutes to options available closer from home, so the mass of restaurants far from home in dense area exert little attraction on a traveler. Two extensions of the model in section 7

[60] This variable, from the NHTS, is equal to the retail price of gasoline as provided by the US Energy Information Administration on the Monday closest to an individual travel day. The prices vary across 5 Petroleum Administration for Defense Districts (East Coast, Gulf Coast, Midwest, Rockies, West Coast (plus Alaska and Hawai)).

[61] Regressions on the model's predicted average trip time (in which all variables in $X_n$ and $Z_n$ should have no effect) also show a negative coefficient, so that most of the effect that I estimate can be due to the correlation of speed with measures of restaurant location (indeed, speed has a positive effect in regressions on trip distance).

relax the IIA property of the logit model. Second, measurement error biases OLS estimates towards zero. Third, omitted variables can bias the coefficient on global density. An instrumental variable strategy alleviates both measurement error and omitted variable biases.

***IV regressions***

If individuals sort into areas based on $\gamma$ or $\sigma$, then OLS coefficients are biased because of a correlation between the error term of equation (11) and measures of restaurant location. For instance, Proposition 4 suggests that individuals with high value of travel time make shorter trips. Therefore, sorting of high value of travel time individuals into dense areas could explain why the model overestimates trip length in areas with high global density. In this case, an instrumented coefficient on global density would be more positive, and closer to the model's prediction. The reverse happens if individuals with marginal preferences or pronounced tastes for variety sort into dense areas; with sorting on $\sigma$ an instrumented coefficient on global density would be even more negative.[62]

An instrument $z_k$ for global density in area $k$ must satisfy two criteria. First, it must be relevant, i.e. correlated with global density conditional on other controls: $corr(global\_density_k, z_k|controls) \neq 0$. Second, the instrument must be exogenous, i.e. uncorrelated with the error term: $corr(\epsilon_{nk}, z_k|controls) = 0$. I instrument global density with the growth in population density from 2000 to 2007, in the county in which an individual lives. Crucially, I am able to select a sample of individuals - old, married, homeowners - very unlikely to move out of county in any given year.

The population data come from the 2000 Census and from the 5-year estimates of the 2009 American Community Survey (the last ACS using the 2000 Census geography). The 5-year estimates provide an average county population count over the years 2005, 2006, 2007, 2008m and 2009, from a 1/8 sample of the population. I use this average as a measure of county population in 2007. County areas are from the Missouri Data Center's MABLE Geocorr2K database, and population density for each county is population count over area. The growth in population density from 2000 to 2007 is the ratio of the log density in 2007 to that in 2000.

[62]Endogenous restaurant supply can also lead to a positive relationship between $\sigma$ and restaurant density. In a model with exogenous, uniform and continuous density of individuals with the same $\sigma$, the density of restaurants in a free-entry equilibrium increases with $\sigma$. I present a richer version of this model is Appendix Appendix D.

While counties vary in size, they are the census geographic units that most closely match an area accessible through 45 minutes of travel.[63] Current county population density is a strong predictor of restaurant global density (table 2). More important, growth in county population density in the 2000s explains variations in the level of restaurant global density in 2011, especially if one controls for initial county population density in 2000.[64] So the instrument is relevant.

The instrument fails the exogeneity requirement if individuals sort into densely populated areas based on characteristics that affect trip time. Using growth to instrument a level is an important step towards satisfying the exogeneity condition. I can control for the initial level of population density in 2000, and people probably seldom choose to reside in an area based on an accurate prediction of its density growth prospect a few years hence. Therefore, the main threat to the exogeneity condition comes from individuals who moved between 2000 and 2007 into areas whose population densities were high because of recent growth. In this case, sorting occurs on the instrument. This is a particular concern because 15.4% of Americans surveyed by the 2009 ACS had changed residence over the previous year, according to Ihrke, Faber, and Koerber (2011). To remedy this, the identification strategy relies on creating a sample of individuals with a low probability of moving out of county in any given year. The moving rate of individuals aged 45 and older is about 7%, with only a 3% chance each year of moving out of county. Homeowners have a 6.7% moving rate, almost five times smaller than that of renters. Married individual also have a lower than average moving rate at 9.9%. Keeping in mind that older individuals are also more likely to be married and to own a home, suppose that 2.5% of married homeowner 55 years and older, that I keep in my sample, randomly move out of county every year.[65] Then more than 80% of these travelers that I observe in 2008 and early 2009 lived in the same county in 2000. That is, the IV regressions are informative, but the results should be treated with caution.

Table 7 contains the two-stage least squares (TSLS) estimation results. The estimates in columns 1–4 are for the full sample, and those in columns 5–8 are for the sample of individuals with a low probability of moving. In each column, the elasticity of trip time with respect to global density

[63]The median county in my sample has a radius of about 38 miles, while 45 minutes of driving usually covers about 25-30 miles in distance.

[64]The instrument is marginally weak if a control for initial population density is not included, for instance in columns 1 and 5 of table 7.

[65]I keep individuals 55 years and older instead of 45 years and older because a 45-year-old in 2008 was 37 in 2000. I observe both age and homeownership status in the NHTS and I select individuals living in households with two or more members to proxy for marital status. This selection still leaves a significant fraction of my NHTS sample of trips (around 40%).

Table 7: The determinants of trip time, with instrument for global density

| Trip time to restaurant | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) ) |
|---|---|---|---|---|---|---|---|---|
| log Global density | $0.102^c$ | $0.089^b$ | $0.083^c$ | $0.120^c$ | $0.239^b$ | $0.200^a$ | $0.161^b$ | $0.230^b$ |
| | (0.055) | (0.043) | (0.047) | (0.063) | (0.102) | (0.072) | (0.069) | (0.10) |
| log Time to closest rest. | $0.269^a$ | $0.229^a$ | $0.213^a$ | $0.261^a$ | $0.353^a$ | $0.249^a$ | $0.224^a$ | $0.271^a$ |
| | (0.039) | (0.019) | (0.017) | (0.023) | (0.076) | (0.029) | (0.025) | (0.037) |
| log Skewness | 0.016 | $0.027^b$ | $0.031^a$ | -0.023 | -0.033 | -0.002 | 0.012 | $-0.059^c$ |
| | (0.021) | (0.013) | (0.012) | (0.02) | (0.038) | (0.021) | (0.015) | (0.031) |
| log Local density | | | | $-0.17^a$ | | | | $-0.210^a$ |
| | | | | (0.022) | | | | (0.036) |
| log Specialization | | | | -0.065 | | | | 0.171 |
| | | | | (0.139) | | | | (0.228) |
| **Controls** | | | | | | | | |
| log County pop. density in '00 | | X | X | X | | X | X | X |
| Individual characteristics | | | X | X | | | X | X |
| Trip characteristics | | | X | X | | | X | X |
| **Instrument** | | | | | | | | |
| $\Delta$ log County pop. density '00-07 | X | X | X | X | X | X | X | X |
| **Samples** | | | | | | | | |
| Full sample | X | X | X | X | | | | |
| Low probability of moving | | | | | X | X | X | X |
| Observations | 11811 | 11811 | 10652 | 10322 | 5520 | 5520 | 4977 | 4815 |
| First-stage stat. | 14 | 33 | 30 | 17 | 11 | 30 | 29 | 14 |

*Notes:* Regressions with a constant in all columns. Robust standard errors, clustered at the county level in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. Dependent variable is trip time to restaurant in all columns. Individual characteristics include 17 dummies for household income, 4 dummies for education, household size, 6 dummies for age, a dummy for gender, a dummy if black and a dummy for worker's status. Trip characteristics include 5 dummies for each peak hour (7-8am, 8-9am, 15-16pm, 16-17pm, 17-18pm), a dummy for whether the trip was followed by a return trip home, a dummy for trips on week-end, the number of children on the trip, the number of adults on the trip, the number of non-household members on the trip, the log time spent at destination, the log speed of a one-mile trip in the driver's census tract and the log of gas price in the driver's region on the week the trip was taken. The sample with a low probability of moving consists of all homeowners 55 years and older living in an household with at least 2 members. The first-stage F statistics are cluster-robust.

is positive and significantly larger than any of the OLS elasticities. This result is consistent with the sorting of individuals with high value of travel time, who are predicted to make shorter trips, into dense areas. The elasticities are about 0.10 in columns 1–4 and are twice as large in columns 5–8, suggesting that sample selection is important for identification. For the IV regressions on the sample of individuals with a low probability of moving, one cannot reject that the effect of global density on trip time is the same as that predicted by the model. The result is robust to removing the control for population density in 2000 (column 5) and to adding controls for individual and trip characteristics (column 7–8). The regression with all controls, in column 8, generates results remarkably similar to those obtained from predicted trip time data in column 5 of table 5. The IV strategy probably mitigates the impact of measurement error, as it moves all coefficients on measures of restaurant location away from 0 and towards the model's prediction.[66]

***Where do welfare gains from density come from?***

Regressions analysis can also distinguish the share of the consumption value of density that is due to gains from variety from the share that is due to savings on travel costs on each trip. To approximate these shares, I run regressions of trip time on 9 dummies for the deciles of the variety-adjusted restaurant price index $P_k$.[67] I exclude the dummy for the first decile, to obtain the average increase in travel time within each upper decile. Within its first decile, the price index is on average equal to 8.68, and within the second decile it equals 9.17 with a round-trip on average 0.46 minutes longer. There is a \$0.49 difference in the two indices and a \$0.092 saving on transport costs, so I attribute 19% of the gains from density to travel time savings for an individual in the second decile moving to the first decile. An individual living in the last decile faces an index equal to 13.23 and makes round-trips 10.6 minutes longer, so 46% of his welfare gains from moving to the first decile are due to shorter trips. The same exercise for other deciles leads to a share of travel costs savings in the gains from density that is generally lower than 50%. Note that IV regressions suggest that part of the reason for shorter trips in dense areas is the sorting of individuals with high value of time, and that extensions of the models (section 7) lead to gains from variety that are slightly larger. I conclude that the gains from density are mostly

[66] The idea that individuals with high value of travel time sort into dense areas is intuitive, but it is only a prediction of the model subject to some caveats. When computing welfare gains, I can confirm that the gains from moving into areas with the highest global density is larger for individuals with higher value of travel time, as is their loss from moving into areas with the lowest global density. In theory however, assuming a uniform density and a fix position for the closest and farthest restaurant, the gains from increasing density are higher for low value of time individuals. The opposite result arises in the data because areas with high global density also have more restaurants very close from home. Clearly, the variables for local density and travel time to the closest restaurant are also endogenous. However, the instrument (even if defined at the census tract instead of at the county-level) is much weaker for these variables, and these IV regressions are sensitive to the set of controls and often lead to unreasonable values. Of course, another explanation for why the instrument works better in large areas is that individuals do not only sort into an area based on how close the nearest few restaurants are, but also take into account the location of many other kinds of destinations.

[67] I use the sample of all round-trips to a restaurant from home by a driver, and include trip and individual characteristics.

gains from being able to visit a destination that one prefers, but that there are also benefits from shorter trips.

*Regression using additional restaurant characteristics from Yelp*

I now use information from Yelp to include additional measures of restaurant characteristics into the regression in equation (11). I only have Yelp data for 23% of restaurants, all in the 20 largest MSAs in my sample. I restrict the sample to trips in areas for which I have Yelp data on at least 50% of restaurants within 45 minutes of travel. This leaves 1531 trips, out of 11884 in the original sample. I compute 7 new variables: mean price, mean ratings, mean number of reviews, mean ease of parking, mean need for reservation, mean attire and mean ambience for all restaurants within 45 minutes of travel. I run a different regression for each new variable.[68] Most of these variables have no significant effect on trip time, except for the mean number of Yelp reviews which has a small positive impact. Mean attire and ambience have a negative, marginally significant effect on trip time (for these two variables, often missing on Yelp, a high mean implies a more upscale environment).

For all 7 of these measures of restaurant characteristics, the standard deviation and the mean in a given area are highly correlated, so including standard deviations in the regression as measures of restaurant diversity leads to similar result. I experiment with using the ratio of a variable's mean for restaurants from 0 to 22.5 minutes of travel to its mean for restaurants from 22.5 to 45 minutes and do not find significant results. Regressions on means computed for the 100 closest restaurants predicts shorter trips if the local area contains expensive, upscale restaurants in which reservations are more likely to be required.

Finally, note that there is a small positive correlation between mean price and global density. However, controlling for price differences across areas did not reconcile the data with the logit model's prediction on the effect of global density. Similarly, estimating a logit model in which meal prices vary by area leads to almost the same result as if meal prices were constant. Average Yelp quality ratings, however, do not vary systematically with density levels, and display little geographic variation.[69] Of course, Yelp ratings often originate from local residents and are not

[68] With this sample, the coefficients on density global is more negative (about -0.05) and that on time to the closest restaurant is smaller than in regressions on the full sample. Areas with a majority of restaurants on Yelp are not randomly selected, and I only collected Yelp data in the largest MSAs, which may differ from the average area.

[69] Berry and Waldfogel (2010) also do not find evidence that average restaurant quality increases with market size.

necessarily comparable across areas. More precise data on price and quality ratings may confirm or infirm these results.

## 7. Extensions of the logit model

In this section, I estimate four extensions of the basic logit model of travel demand. First, I allow sorting into dense areas by value of travel time, in line with the IV regression results. Second, I let meal prices vary with travel time from home, to deal with a potential source of bias in my estimates. The last two extensions relax the IIA property of the logit model and introduce restaurant diversity into the model. When the IIA does not hold, restaurants become redundant and many of the restaurants far from home in dense areas exert little attraction on an individual, because they are similar to restaurants available closer to home. This could explain why the OLS effect of global density on trip time is smaller than what the logit model predicts.

### *Sorting by value of time*

IV regressions indicate that individuals who choose to live in high restaurant density areas make shorter trips, and Proposition 4 suggests that sorting by value of travel time can explain that result. Hence, I estimate an extension of the basic logit model in which an individual's value of travel time $\gamma$ varies with global density in area $k$. Denote the value of travel time for the driver of trip $n$ in area $k$ by $\gamma_{nk}$, with $\bar{\gamma}$ the geometric mean of $\gamma$ over all $n$. To be consistent with previous estimation, I set $\bar{\gamma}$ = 0.2. I parameterize $\gamma_{nk}$ as:

$$\gamma_{nk}(\beta,T_k) = \bar{\gamma}\left(\frac{\overline{global\_density}}{global\_density_k}\right)^{-\beta}, \tag{12}$$

where $\overline{global\_density}$ is the geometric mean of global density. $\beta$ captures the strength of the relationship between $\gamma_{nk}$ and global density. If $\beta > 0$, then individuals living at above average density levels have higher values of travel time. One can use equation (12) to verify that $\bar{\gamma}$ is indeed the geometric mean of $\gamma_{nk}$, i.e. that $\bar{\gamma} = (\prod_{n=1}^{N}\gamma_{nk})^{1/N}$.[70] As in section 4, the probability of a trip to a restaurant at travel time $m_{nk}$ is:

$$prob(m_{nk}|\sigma,\beta,T_k) = \frac{(p+2\gamma_{nk}(\beta,T_k)m_{nk})^{-\sigma}}{\sum_{i=1}^{I_k}(p+2\gamma_{nk}(\beta,T_k)t_{ki})^{-\sigma}},$$

[70] The geometric mean is always smaller than the arithmetic mean, but in this case the arithmetic mean of $\gamma_{nk}$ that I obtain after estimating the model is almost identical to $\bar{\gamma}$.

the log-likelihood function is:

$$\ell(\sigma,\beta,m,T) = \sum_{n=1}^{N} log(prob(m_{nk}|\sigma,\beta,T_k)),$$

and the maximum likelihood estimator is:

$$(\hat{\sigma},\hat{\beta}) = \underset{\sigma,\beta}{\operatorname{argmax}}\, \ell(\sigma,\beta,m,T). \tag{13}$$

I find $\hat{\sigma}$ =10.1 and $\hat{\beta}$ = 0.21 (in column 2 of table 3). A plot of the log-likelihood function suggests that it is concave for any reasonable values of the parameters and a likelihood ratio test easily rejects the model without sorting in favor of the model with sorting by value of time. Recall that the model without sorting predicts the IV regression results of a positive effect of global density. Consistent with this, re-estimating equation (10) for the model with sorting shows that it correctly predicts the OLS regression results of a near zero effect of global density on trip time.

The elasticity of substitution $\sigma$ is slightly lower in the model with sorting, because the low share of trips to remote restaurants in dense areas is explained by the positive relationship between global density and value of time, not by higher substitutability between restaurants. This lower elasticity of substitution implies marginally higher gains from variety in the model with sorting, for instance the aggregate gains from being able to visit restaurants other than the closest one increase by 9%.

$\hat{\beta}$ = 0.21 implies large but not implausible variation in value of travel time. For instance the $10^{th}$ percentile of $\gamma_{nk}$ is at $\gamma$ = 0.14 and the $90^{th}$ percentile is at $\gamma$ = 0.28. This estimate of $\beta$ is almost certainly biased upward, because $\gamma$ also captures variation in trip time that may be attributable to restaurants becoming redundant, or to measurement error. I am not aware of other value of travel time estimates that vary with measures of restaurant density in the home area. Abrantes and Wardman (2011), in a meta-analysis of 140 recent studies for the United Kingdom, run regressions of value of travel time estimates on various explanatory variables. They find that value of travel time is 27% larger in London and the South East of England (which is more densely populated) but do not uncover any other spatial or density effects. However, they find strong evidence that value of travel time is higher, by 70%, in congested traffic. Given that car travel speed is lower in dense areas, congestion aversion could also explain, without any sorting, why a model with higher value of travel time in dense areas fits the data better.

*Letting average meal price vary with travel time from home*

Estimates of the elasticity of substitution between restaurants are biased if the characteristics of restaurants close to home differ systematically from those of restaurants far from home.

I use data from Yelp to compute average characteristics of restaurants that vary with travel time. I place each restaurant in the choice set of a traveler in a five minutes time bin, from 0–5 to 40–45 minutes. There are not enough restaurants on Yelp in each time bin to compute measures that vary both by travel time and by areas, so I compute an average that varies only with travel time. To do so I create a set of all restaurants 0–5 minutes away from any individual. Then I compute the mean of Yelp variables over all restaurants in this set, with some restaurants double-counted if they are within 5 minutes of travel from more than one individual. I repeat the process for restaurants 5–10 minutes away, and so on. Locations on average far from home have the same average quality ratings, but are pricier, have more reviews, a more upscale ambience and attire, and are more likely to require a reservation. The average price of a restaurant within 0–5 minutes of travel from home is \$12.56, which increases to \$13.33 for restaurants 40–45 minutes away.[71] Parking is also harder in restaurants far from home; 22% of restaurants within 40–45 minutes of travel have street parking only, up from 15% of restaurants 0–5 minutes from home. One interpretation of these results is that a majority of travelers live in medium density suburbs, relatively far from restaurants in high-density areas which tend to be slightly more expensive and upscale, and less likely to have a private lot for parking.

To assess the sensitivity of my results to these spatial variations in restaurant characteristics, I re-estimate the logit model, but with meal price in each time bin equal to its average value. For instance, I set the price of each restaurant 0–5 minutes from home at \$12.56. Given that restaurants farther from home are not only more expensive (making them less attractive), but also more upscale (making them more attractive), an estimate in which only meal price varies is a lower bound for the downward bias on $\sigma$ caused by the correlation between meal price and travel time. I find $\sigma = 10.0$ (column 3 of table 3). This elasticity of substitution is lower than that in the model with constant meal prices, because to some extent the lower share of trips to restaurants

[71] I also find that restaurants far from home are less likely to be part of restaurant chains, and the next extension takes it into account. The specialization index also does not vary with travel time, but my methodology is imperfect in this case, as it could be that restaurants close to any given individual are usually of the same category, but that categories differ enough across areas to lead to a low average specialization. The nested-logit extension takes the exact category of each restaurant into account.

farther away is explained by price differences, not by high substitutability. In turn this lower elasticity leads to aggregate gains from restaurant variety about 5% larger.

### *Redundant chain restaurants*

According to a census of all commercial restaurants conducted by the NPD group in the Fall 2010, there are 267,499 chain restaurants in the US, representing 46% of the total.[72] Two restaurants in the same chain are never exactly the same, but it must be true that two McDonald's are highly substitutable with one another, so a model in which restaurants in the same chain are perfectly substitutable may be a better representation of reality. This is perhaps the simplest way to relax the logit assumption that all restaurants are equally substitutable (IIA property) and to introduce restaurant diversity into the model. Intuitively, areas consisting mostly of repeated chain restaurants have low diversity, and the model is now flexible enough to take this into account. A model with perfectly substitutable chains could also predict the small effect of global density on trip time better than the logit model, because some of the thousands of restaurants far away from home in dense areas necessarily belong to chains available closer from home, and as such exert no attraction on a traveler.

To estimate this model, I code the 50 largest restaurant chains in my data, which represent 23% of all restaurants in the sample, and are likely to be occur more than once within 45 minutes of travel. Because travel is costly, a restaurant that is perfectly substitutable with another restaurant closer from home is never visited. To understand what this implies, suppose that an individual's choice set in the logit model is $T_k = \{t_{k1}, \ldots, t_{kI_k}\}$. If restaurants located at $t_{k5}$, $t_{k9}$ and $t_{k104}$ are all Subway restaurants, then the choice set if chains are perfectly substitutable becomes $T'_k = \{t_{k1}, \ldots, t_{kI_k}\} \setminus \{t_{k9}, t_{k104}\}$. Similarly, I eliminate all repeat restaurants in the other 49 largest chains. Estimation is then exactly as in the logit model, with the estimator given by equation (8), but with a choice set $T'_k$ containing no repeat chain restaurants. I find $\hat{\sigma} = 10.0$ (column 4 of table 3).

In an alternative specification, I assume that restaurants with the same Google Places title are perfectly substitutable. This likely eliminates almost all repeat chain restaurants (54% of the restaurants in my sample are part of a group of two or more restaurants with the same title) but leads to more errors, as some restaurants with the same title are not part of the same chain, and

[72]See http://www.npd.com/press/releases/press_110124.html for the press release.

many restaurant chains have entries on Google Places with inaccurately spelled or shortened title. In this case, I find $\hat{\sigma}$ = 9.7 (column 5 of table 3)

The elasticity of substitution is lower when restaurants are perfectly substitutable within chains. The reason is, the logit model attributes the low share of trips to remote restaurants to high substitutability between restaurants in general, instead of just assuming that some of these restaurants are perfectly substitutable with eating options closer from home. The two extensions above generate predictions that are only slightly closer to the data; when I eliminate repeat titles I obtain a predicted effect of global density on trip time one percentage point closer to its actual effect. The aggregate gains from variety are only marginally larger than those from the model without redundant chains.[73] Perfect substitutability within chains barely changes the welfare estimates because the percentage of restaurants that are part of a chain is larger in low-density areas. Contrary to an intuition that would be correct if restaurants were randomly distributed, repeat restaurants are less common in dense areas.

***Nested-logit model***

In a nested-logit model, individuals first choose a category of restaurants (e.g. Pizza, Chinese or Burger) and then decide which restaurant to visit within that category. The IIA property of the logit model does not hold, because restaurants within the same category are more substitutable. To solve the demand-side of the model, I adapt the strategy in Sheu (2011), who proves that a nested-logit model can generate the same choice probabilities as a nested-CES model. To recover unobserved taste parameters for each category of restaurants in each area, I derive an equilibrium condition on the supply-side, assuming zero profits and free-entry of restaurants.

There are 85 categories of restaurants, indexed by $c$.[74] There is an exogenous distribution of tastes for restaurant categories that is specific to each area $k$, and a parameter $b_{kc}$ captures the

[73]The flip side of this result is that in reduced-form regressions, controls for the percentage of chain restaurants (or the percentage of repeat chains or repeat titles) do not have a significant effect on trip time and do not affect the coefficient on global density.

[74]The categories, and their percentage share, are listed in Appendix Appendix E. The category is 'undefined' for 17% of restaurants, usually smaller independent places serving standard fares. I code all other restaurants into 84 categories (using information from Google Places), that closely match the categories that Yelp distinguishes on its website.

taste for category $c$ in area $k$.[75] Let $i$ index the number $I_{kc}$ of restaurants in category $c$ within 45 minutes of travel in area $k$. Travel time to restaurant $i$ in category $c$ in area $k$ is $t_{kci}$ minutes. Meal price is constant and the same within each category, so the total price of eating at restaurant $i$ is $p_{kci} = p + 2\gamma t_{kci}$.[76] As before, each restaurant receives a type I extreme value idiosyncratic shock $\epsilon_{kci}$ with scale parameter $1/(\sigma - 1)$, but now each restaurant also receives a category-specific type I extreme value idiosyncratic shock $\varsigma_{kc}$ with scale parameter $1/(1-\mu)$. An individual's budget constraint is $p_{kci} r_{kci} = y$, where $y_k$ is expenditures on restaurants in area $k$, and $r_{kci}$ is the number of trips to the chosen restaurant $i$. The utility from choosing restaurant $i$ from category $c$ in area $k$ is:

$$u_{kci} = ln\left(b_{kc}^{1/(\sigma-1)} r_{kci}\right) + \epsilon_{kci} + \varsigma_{kc},$$

and the indirect utility is:

$$v_{kci} = ln(y_k) + \frac{1}{\sigma - 1} ln(b_{kc}) - ln(p_{kci}) + \epsilon_{kci} + \varsigma_{kc}.$$

The traveler first solves the optimization problem within each restaurant category, exactly as in section 2, and then chooses the category that maximizes his expected utility. The probability of a trip of length $t_{kci}$, given the set of travel times $T_k = \{T_{k1}, \ldots, T_{kc}, \ldots, T_{k85}\}$, where $T_{kc} = \{t_{kc1}, \ldots, t_{kci}, \ldots, t_{kcI_{kc}}\}$, is:

$$prob(t_{kci}|T_k) = \frac{b_{kc} P_{kc} p_{kci}^{-\sigma}}{\sum_{c=1}^{85} b_{kc} P_{kc}^{\sigma-\mu} \sum_{i=1}^{I_{kc}} p_{kci}^{-\sigma}} \tag{14}$$

where $P_{kc} = \left(\sum_{i=1}^{I_{kc}} p_{kci}^{1-\sigma}\right)^{1/(1-\sigma)}$ (see Sheu (2011) for a proof). It is easy to show that $\sigma$ is an elasticity of substitution within categories and that $\mu$ is an elasticity of substitution across categories. One expects $\sigma > \mu$ if an individual cares more about restaurant category than about the particular restaurant that she visits within a category (e.g. she wants to eat Sushi, regardless of the exact restaurant).

Equation (14) is not suitable for estimation because area-specific taste parameters are unobserved.[77] If these tastes for categories are exogenous, in equilibrium they should affect the share

[75] Schiff (2012) shows that even if preferences are identical everywhere, in a free-entry model densely populated areas feature more categories of restaurants because they contain enough people with marginal tastes to make restaurants of the least popular categories profitable. This argument is intuitive, but it alone cannot account for the wide range of restaurant diversity that I measure in areas at the same density level. For instance, some areas in Texas contain a vast majority of Mexican restaurants, which likely reflects a special taste for this category in these areas.

[76] Constant meal price across categories is an equilibrium result that depends on assuming constant meal costs.

[77] Estimating the model without taste parameters for categories means that marginal and arbitrarily defined restaurant categories (e.g. a 'soup' category is defined on Yelp) exert an implausibly large attraction on travelers.

of restaurants of each category in an area. To obtain an expression of this relationship, I introduce restaurant supply into the model. The details are in Appendix Appendix D (I assume free-entry and use a continuous version of the model with uniform density to obtain an analytical result). Denote by $d_{kc}$ the uniform density of restaurants of type $c$ in area $k$. In equilibrium, for any two categories $c$ and $c'$:

$$\frac{d_{kc}}{d_{kc'}} = \left(\frac{b_{kc}}{b_{kc'}}\right)^{\frac{\sigma-1}{\sigma-\mu}}. \tag{15}$$

The relative density of restaurants in these two categories depends on relative tastes in an intuitive way. If individuals care little about variety in restaurant categories (i.e. if $\sigma \approx \mu$), then in equilibrium there are infinitely more restaurants of the preferred restaurant category. As $\sigma$ becomes larger than $\mu$, relative restaurant density tends to a reflection of relative tastes.

I now estimate the parameters $\sigma$ and $\mu$. Let $n$ index each trip in the sample. The choice set of a traveler in area $k$ is $T_k$. I compute the density ratio in equation (15) as a ratio of the number of restaurants within 45 minutes of travel in each category:

$$d_{kc}/d_{kc'} = I_{kc}/I_{kc'}. \tag{16}$$

After normalizing $b_{k1} = 1$, I plug equation (16) into equation (15) to obtain an expression for each unobservable taste parameter $b_{kc}$ as a function of $I_{kc}$ and $I_{k1}$, which are observable. Substituting $b_{kc}$ into equation (14) results in an expression for the probability of visiting each restaurant that depends only on the parameters $\sigma$ and $\mu$, and on the restaurant choice set $T_k$:

$$prob(t_{kci}|\sigma,\mu,T_k) = \frac{\left(\frac{I_{kc}}{I_{k1}}\right)^{\frac{\sigma-\mu}{\sigma-1}} P_{kc}(p+2\gamma t_{kci})^{-\sigma}}{\sum_{c=1}^{85}\left(\frac{I_{kc}}{I_{k1}}\right)^{\frac{\sigma-\mu}{\sigma-1}} P_{kc}^{\sigma-\mu}\sum_{i=1}^{I_{kc}}(p+2\gamma t_{kci})^{-\sigma}}. \tag{17}$$

Without data on the exact restaurant choice of a traveler, equation (17) is less amenable to estimation by maximum likelihood than the equivalent equation from the logit model, because travel time is no longer the only source of observed heterogeneity among restaurants. Denote travel time for trip $n$ by $m_{nk}$. To estimate the model, let $R_{nk}(m_{nk})$ be a set of restaurants at travel time 'close' to actual trip time $m_{nk}$. With a slight abuse of notation, let $i$ index all restaurants in $R_{nk}$, so the log-likelihood function is:

$$\ell(\sigma,\mu,m,T) = \sum_{n=1}^{N} log\left(\sum_{i\in R_{nk}(m_{nk})} prob(m_{nk}|\sigma,\mu,T_k)\right),$$

and the maximum likelihood estimator is:

$$(\hat{\sigma},\hat{\mu}) = \underset{\sigma,\mu}{\text{argmax}}\ \ell(\sigma,\mu,m,T). \tag{18}$$

The estimates are sensitive to the exact definition of the set $R_{kn}$, in particular as it affects the probability weight that the estimator places on the shortest trips. As an example, define $R_{nk}$ as the set of all restaurants within 5 minutes of $m_n$. In this case, if the closest restaurant is 9 minutes from home and $m_{nk}$ = 30, then $R_{kn}$ is the set of all restaurants from 25 to 35 minutes of travel. Now suppose that $m_{nk}$ = 10. If I use a rule forcing $R_{kn}$ to contain 10 minutes of restaurants – in the example $R_{kn}$ would then be the set of all restaurants from 9 to 19 minutes – I find $\hat{\sigma}$ = 11.3 and $\hat{\mu}$ = 2.2. However, if I use another rule such that the 10 minute time bin is always symmetric around $m_{kn}$ – in the example $R_{nk}$ would be the set of all restaurants between 5 and 15 minutes – I find $\hat{\sigma}$ = 11.2 and $\hat{\mu}$ = 5.2.[78] I conclude that I cannot, using only data on the distribution of trip times, provide robust estimates of the nested-logit's parameters.

If the difference between $\sigma$ and $\mu$ is large, then additional restaurants rapidly become redundant. There is little need for long trips in dense areas because travelers care mostly for restaurant categories, of which there are relatively few. In regressions on the model's predicted average trip time, the effect of global density on trip time is about 50% closer to that observed in the data if $\hat{\sigma}$ = 11.3 and $\hat{\mu}$ = 2.2, and 20% closer if $\hat{\sigma}$ = 11.2 and $\hat{\mu}$ = 5.2.

Unlike the logit model, the nested-logit predicts the effect of restaurant diversity on travel. In a nested-logit model, the accumulation of restaurants far from home is unattractive if it consists of restaurants in repeated categories. If all restaurants in an area are in different categories, then predicted trip times are large because the elasticity of substitution between restaurants is low (equal to $\mu$, the elasticity across categories). The reverse is true if all restaurants are in the same category, because in this case the elasticity of substitution, $\sigma$, is high. This important feature of the model is consistent with the data, although the effect of the specialization index on trip time

[78] This second rule places more weight on short trips, which increases the value of $\mu$ towards that of $\sigma$. The reason is, if $\mu = \sigma$ the model predicts that the closest restaurant is visited with highest probability, while if $\mu < \sigma$ individuals may visit with highest probability a restaurant from a preferred category, not necessarily the closest. Because of measurement error on travel time to the closest restaurant, I have to assume that whenever travel time to the closest restaurant is lower than trip time, the two are equal. This arbitrary assumption means that there are many trips to exactly the closest restaurant in my sample, which provide false evidence that $\mu = \sigma$. For this reason, it is preferable to define $R_{nk}$ in a way that does not put too much probability weight on the shortest trips, and I have more confidence in the estimates $\hat{\sigma}$ = 11.2 and $\hat{\mu}$ = 2.2.

predicted by the nested-logit model is smaller than that observed in the data.[79]

I also compute a variety-adjusted restaurant price index from the nested-logit model. Using $\hat{\sigma}$ = 11.2 and $\hat{\mu}$ = 5.2, I obtain values of the price index close to those in the logit model. Performing the last two counterfactual experiments of section 5 with $\hat{\sigma}$ = 11.3 and $\hat{\mu}$ = 2.3, I find an average relative price index of $P_{k',k}$ = 1.69 with counterfactuals at the $95^{th}$ percentile of global density, and of $P_{k',k}$ = 0.77 with counterfactuals at the $5^{th}$ percentile.[80] Hence, spatial variation in diversity can exacerbate welfare differences across areas if individuals care much about access to a variety of categories. Only further research, ideally with new data on restaurant choice, can refine this finding.

## 8. Discussion: Accessibility

The variety-adjusted price index $P_k$ is also a travel accessibility index. The index has the advantage of great geographical coverage, ease of interpretation, and a basis in precise microgeographic data, in the context of restaurant accessibility. It allows for variation in travel speed across areas, and places less weight on remote locations. I therefore use it to provide tentative numerical answers to some fundamental questions on the links between transportation technology and urban form, and on the trade-offs between different spatial organizations of society. I present numbers for the basic logit model.

As already mentioned, the median value of the index, at about 10, is significantly lower than the average price of a restaurant meal ($13) before including transport costs. This result supports the argument in Glaeser and Kahn (2004), that the suburban lifestyle shared by a majority of Americans offers good accessibility through fast car travel. These authors also argue that the

[79] The predicted effect is small because the number of categories is dwarfed by the number of restaurants available. The difference bewteen the actual and predicted effect could also arise because specialization captures something else in the data that is not controlled for. Or individual travelers may, on each trip, only make a decision based on, or be aware of, a subset of all available restaurants. Note however that when computing welfare gains, the mass of restaurants beyond 30 minutes of travel has essentially no impact on welfare and that there is no need for an individual to have perfect information on thousands of remote restaurants for my estimates to be valid. It is easy to estimate a version of the logit model in which restaurants farther away from an individual are known (i.e. part of his choice set) with smaller probability (which is parameterized such that 100% of restaurants at 0 minute from home are known, and the probability of knowing restaurants farther away decreases with travel time). The estimation results are not precise but in the model that best fits the data, individuals know only about 69% of restaurants 45 minutes away. This is not large enough to have any effect on estimated welfare gains. I also estimate, by simulated maximum likelihood, a model in which the scale of the type I extreme value distribution of the error term decreases with distance, and obtain similar results.

[80] To make comparisons easier, I assume that the distribution of restaurant categories in the counterfactual area depends on the taste for categories of the individual forced to move. Also, I cannot compute price indices for the experiment in which the counterfactual area consists only of the restaurant closest to home, because it violates the assumption that the distribution of restaurant categories is a reflection of tastes for categories.

automobile is the leading cause of urban sprawl in the United States. While it is difficult to establish a causal relationship - for instance Meyer, Kain, and Wohl (1965) argue that decentralization predates the automobile - undeniably the current spatial configuration of Americans hardly makes sense without rapid private transportation.[81] Yet most Americans do have at least one restaurant within walking distance; the median walking time to the closest restaurant is about 11 minutes, with a much larger mean at 22 minutes (the corresponding numbers by car are 4.2 and 5.3 minutes). The gains from variety when traveling on foot, however, are considerably smaller. If travelers have to walk, the median value of the index increases to 14.5, with a mean at about 18.[82] There are strong complementarities between car transportation and suburban living, but in high-density urban areas with slow car travel, the price index when traveling by car and by foot are much closer. This explains why 65% of Manhattanites walk to a restaurant.

In fact, the lowest price index, or equivalently the highest level of accessibility, belongs to areas with the *slowest* car travel. Faster travel speed mechanically decreases the price index, but the correlation between speed and the index is still positive (the rank correlation is 0.22, but would certainly be much stronger without measurement error on speed). Raising the density of destinations, of population and of the street network reduces travel speed, but not enough to annihilate the benefits from greater access to destinations. Indeed, a few areas in New York City have both the lowest car travel speed and the lowest price index. A walker in these high-density areas faces a lower price index than car driver in 95% of all block groups.

## 9. Conclusion

This paper shows how to estimate the consumption value of density by combining travel data with microgeographic data on local businesses. Individuals' substitution patterns among travel destinations reveal gains from urban density that are large but localized. These gains originate in part from shorter trip times, but mostly arise because increased choice in denser areas allows individuals to visit destinations that they prefer. The consumption benefits of density that I estimate in the restaurant industry demonstrate that cities, and downtown cores in particular, enjoy a sizable advantage in non-tradable service provision. Using a similar methodology, one

[81] Public transportation is not an option in many suburbs, and its use to eat out is negligible.

[82] These numbers depend on a relatively brisk walking pace of 15 minutes per mile, and I do not account for higher value of travel time by foot (preference for cars), which makes walking an even less attractive option. I also do not account for the price of downtown parking fees, which increases the costs of driving into some of the highest density areas.

could expand the scope of the variety-adjusted restaurant price index to other trip purposes in the NHTS, for instance to trips for medical or dental care. Ultimately, the canonical model of spatial equilibrium (Rosen, 1979, Roback, 1982) implies that positive amenities translate into higher land rents. Investigating the capitalization of greater access to goods and services into real estate prices, at fine spatial scales, could be a productive area for future research.

## References


Abrantes, Pedro A.L. and Mark R. Wardman. 2011. Meta-analysis of uk values of travel time: An update. *Transportation Research Part A* 45(1):1–17.

Ackerberg, Daniel A. and Marc Rysman. 2005. Unobserved product differentiation in discrete choice models: Estimating price elasticities and welfare effects. *RAND Journal of Economics* 36(4):771–778.

Albouy, David and Bert Lue. 2011. Driving to opportunity: Local wages, commuting, and sub-metropolitan quality of life. University of Michigan.

Anderson, Simon P., Andre de Palma, and Jacques-Francois Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. Cambridge MA: MIT Press.

Ben-Akiva, Moshe and Steven Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge MA: The MIT Press.

Berry, Steven and Joel Waldfogel. 2010. Product quality and market size. *Journal of Industrial Economics* 58(1):1–31.

Bhat, Chandra, Susan Handy, Kara Kockelman, Hani Mahmassani, Qinglin Chen, and Lisa Weston. 2000. Development of an urban accessibility index: Literature review. Technical Report Report TX-01/7-4938-1, Center for Transportation Research, The University of Texas at Austin, prepared for the Texas Department of Transportation.

Broda, Christian and David E. Weinstein. 2006. Globalization and the gains from variety. *Quarterly Journal of Economics* 121(2):541–585.

Broda, Christian and David E. Weinstein. 2010. Product creation and destruction: Evidence and price implications. *American Economic Review* 100(3):691–723.

Carlino, Gerald A. and Albert Saiz. 2008. Beautiful city: Leisure amenities and urban growth. Technical report, FRB of Philadelphia Working Paper No. 0822.

Clark, Terry Nichols. 2003. Lakes, opera and juice bars: Do they drive development? In Terry Nichols Clark (ed.) *The City as an Entertainment Machine*, volume 9 of *Research in Urban Policy*. New York NY: JAI Press/Elsevier, 103–140.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2011. The identification of agglomeration economies. *Journal of Economic Geography* 11(2):253–266.

Couture, Victor, Gilles Duranton, and Matthew A. Turner. 2012. Speed. University of Toronto.

Dixit, Avinash K. and Joseph E. Stiglitz. 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67(3):297–308.

Ewing, Reid and Robert Cervero. 2010. Travel and the built environment. *Journal of The American Planning Association* 76(3):265–294.

Feenstra, Robert. 1994. New product varieties and the measurement of international prices. *American Economic Review* 84(1):157–177.

Glaeser, Edward L. and Matthew E. Kahn. 2004. Sprawl and urban growth. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2481–2527.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. 2001. Consumer city. *Journal of Economic Geography* 1(1):27–50.

Handbury, Jessie. 2012. Are poor cities cheap for everyone? Non-homotheticity and the cost of living across US cities. University of Pennsylvania.

Handbury, Jessie and David E. Weinstein. 2012. Is new economic geography right? Evidence from price data. Technical report, NBER Working Paper No. 17067.

Ihrke, David K., Carol S. Faber, and William K. Koerber. 2011. Geographical mobility: 2008 to 2009. Technical report, U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau.

Kockelman, Kara Maria. 2001. A model for time- and budget-constrained activity demand analysis. *Transportation Research Part B-Methodological* 35(3):255–269.

Li, Nicholas. 2012. An Engel curve of variety. University of Toronto.

Melo, Patricia C., Daniel J. Graham, and Robert B. Noland. 2009. A meta-analysis of estimates of urban agglomeration economies. *Regional Science And Urban Economics* 39(3):332–342.

Meyer, John R., John F. Kain, and Martin Wohl. 1965. *The urban transportation problem.* Cambridge MA: Harvard University Press.

Niemeier, Debbie A. 1997. Accessibility: an evaluation using consumer welfare. *Transportation* 24(4):377–396.

Okrent, Abigail M. and Julian M. Alston. 2010. Demand for food in the United States. Technical report, RMI-CWE Working Paper number 1002.

Okrent, Abigail M. and Julian M. Alston. 2011. The demand for disaggregated food-away-from-home products. US Department of Agriculture and University of California, Davis.

Parry, Ian W. H., Margaret Walls, 2007. "Automobile Externalities Winston Harrington, and American Economic Association vol. 45(2) pages 373-399 June. Policies, " Journal of Economic Literature. 2007. Automobile externalities and policies. *Journal of Economic Literature* 45(2):373–399.

Roback, Jennifer. 1982. Wages, rents and the quality of life. *Journal of Political Economy* 6:1257–1278.

Rosen, Sherwin. 1979. Wages-based indexes of urban quality of life. In P. Mieszkowski & M. Straszheim (ed.) *Current Issues in Urban Economics*. Baltimore: John Hopkins Univ. Press.

Sato, Kazuo. 1976. The ideal log-change index number. *Review of Economics and Statistics* 58(2):223–228.

Schiff, Nathan. 2012. Cities and product variety. University of British Columbia.

Sheu, Gloria. 2011. Price, quality, and variety: measuring the gains from trade in differentiated products. US Department of Justice.

Small, Kenneth A. and Erik T. Verhoef. 2007. *The Economics of Urban Transportation*. New York (NY): Routledge.

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation*. Second edition. Cambridge (MA): Cambridge University Press.

U.S. Environmental Protection Agency. 2012. Inventory of US greenhouse gas emissions and sinks: 1990-2010. Technical report, EPA 430-R-12-001.

Vartia, Yrjo. 1976. Ideal log-change index numbers. *Scandinavian Journal of Statistics* 3(3):121–126.

## Appendix A. Proofs

### *Proof of proposition 1*

Recall that:

$$\bar{t}_k = \frac{\sum_{i=1}^{I_k} (p + 2\gamma t_{ki})^{-\sigma} t_{ki}}{\sum_{i=1}^{I_k} (p + 2\gamma t_{ki})^{-\sigma}}.$$

I drop the subscript $k$ for simplicity. Denote the numerator of $\bar{t}$ by $y$ and its denominator by $x$, so that $\bar{t} = y/x$. Suppose that we add a restaurant at exactly $\bar{t}$, with travel time $t_i = y/x$. To show that $\bar{t}$ doesn't change, one needs to show that:

$$\frac{y}{x} = \frac{y + (p + 2\gamma t_i)^{-\sigma} t_i}{x + (p + 2\gamma t_i)^{-\sigma}}.$$

This equation reduces to $t_i = y/x$, which is true by assumption. To prove the second part of the proposition, that adding a restaurant at travel time $t_i > y/x$ increases $\bar{t}$ just note that:

$$\frac{y}{x} < \frac{y + (p + 2\gamma t_i)^{-\sigma} t_i}{x + (p + 2\gamma t_i)^{-\sigma}},$$

whenever $t_i > y/x$. The reverse inequality is true assuming $t_i < y/x$.

### *Proof of proposition 2*

Let density be uniform and $t$ be travel time between each restaurant. The total number of restaurants in an area is equal to $I = M/t$, where $M$ is the number of minutes required to travel through the entire area (e.g. a country). Travel time to restaurant $i$ is $t_i = ti$.

To prove the proposition that $lim_{t \to 0+} \bar{t} > 0$, I show that the total probability of visiting restaurants indexed from $I/2$ to $I$ stays larger than 0 as $t \to 0+$. The probability of a trip to restaurant $i$ is:

$$prob_i = \frac{(p + 2\gamma ti)^{-\sigma}}{\sum_{i=1}^{M/t} (p + 2\gamma ti)^{-\sigma}},$$

so the sum of the probabilities of visiting restaurants indexed from $I/2$ to $I$ is:

$$\frac{\sum_{i=1}^{M/2t} (p + 2\gamma (t(i + I/2)))^{-\sigma}}{\sum_{i=1}^{M/t} (p + 2\gamma (ti)^{-\sigma}} = \frac{\sum_{i=1}^{M/2t} (p + 2\gamma (ti + M/2))^{-\sigma}}{\sum_{i=1}^{M/t} (p + 2\gamma ti)^{-\sigma}}, \tag{A1}$$

where I substituted for $I = M/t$. Equation (A1) becomes a definite integral as $t \to 0+$. To see this, define $x_i = ti$ so that $\Delta x = t$ and multiply both the numerator and denominator of equation (A1) by $\Delta x = t$, to obtain the following (lower) Riemann sum:

$$\frac{\sum_{i=1}^{M/2\Delta x}(p+2\gamma(x_i+M/2))^{-\sigma}\Delta x}{\sum_{i=1}^{M/\Delta x}(p+2\gamma x_i)^{-\sigma}\Delta x}.$$

By assumption, $p > 0$, $\gamma > 0$, $x_i > 0$ and $\sigma > 1$, so each element $(p+2\gamma x_i)^{-\sigma}$ or $(p+2\gamma(x_i + M/2))^{-\sigma}$ is bounded and continuous, and therefore:

$$\lim_{\Delta x \to 0+} \frac{\sum_{i=1}^{M/2\Delta x}(p+2\gamma(x_i+M/2))^{-\sigma}\Delta x}{\sum_{i=1}^{M/\Delta x}(p+2\gamma x_i)^{-\sigma}\Delta x} = \frac{\int_0^{M/2}(p+2\gamma(x+M/2))^{-\sigma}dx}{\int_0^{M}(p+2\gamma x)^{-\sigma}dx}.$$

Note that taking a limit as $\Delta x$ tends to $0+$ is equivalent to taking a limit as $t$ tends to $0+$. Both integrals in the previous equation are easy to solve and lead to finite results, showing that the probability of visiting restaurants far from home is larger than 0 as $t \to 0+$. I conclude that $lim_{t\to 0+}\bar{t} > 0$, which completes the proof.[83]

### *Proof of proposition 3*

As shown in the text, as $\sigma$ tends to infinity the probability of visiting the closest restaurant becomes infinitely larger than that of visiting any other restaurants, so $\bar{t}$ becomes equal to travel time to the closest restaurant. Therefore, adding $x$ minutes of travel to each restaurant (and in particular to the closest one) increases expected trip time by $x$ minutes. $\bar{t}$ is continuous with respect to $\sigma$ (which is always larger than 1), so this result must hold approximately for $\sigma$ large enough.

### *Proof of proposition 4*

To prove that $\frac{\partial \bar{t}}{\partial \sigma} < 0$, I show that as $\sigma$ increases, for any two pairs of restaurants there is a decrease in the ratio of the probability of visiting the restaurant that is farther over the probability of visiting the restaurant that is closer. That is, for any given $t_i > t_j$, it is true that:

$$\frac{\partial\left(\frac{prob_i}{prob_j}\right)}{\partial\sigma} = \frac{\partial\left(\frac{p+2\gamma t_i}{p+2\gamma t_j}\right)^{-\sigma}}{\partial\sigma} = -\left(\frac{p+2\gamma t_i}{p+2\gamma t_j}\right)^{-\sigma} ln\left(\frac{p+2\gamma t_i}{p+2\gamma t_j}\right) < 0.$$

[83]As an aside, note that $\bar{t}$ is a weighted average of all $t_i$, with the weight (probability of a trip) decreasing with $t_i$. So it is immediate that $\bar{t} < t_I$ for all $t > 0$, as the probability of visiting any other restaurant is larger than that of visiting restaurant $I$. Also, it is easy to prove that if the closest restaurant is assumed to be located at $t_1$, then the proposition becomes $lim_{t\to 0+}\bar{t} > t_1$.

Similarly, it must be true that $\frac{\partial \bar{t}}{\partial \gamma} < 0$, because for any $t_i > t_j$:

$$\frac{\partial \left(\frac{p+2\gamma t_i}{p+2\gamma t_j}\right)^{-\sigma}}{\partial \gamma} = -\sigma \left(\frac{p+2\gamma t_i}{p+2\gamma t_j}\right)^{-\sigma-1} \left(\frac{2t_i(p+2\gamma t_j) - 2t_j(p+2\gamma t_i)}{(p+2\gamma t_j)^2}\right) < 0.$$

Finally, it must be true that $\frac{\partial \bar{t}}{\partial p} > 0$, because for any $t_i > t_j$:

$$\frac{\partial \left(\frac{p+2\gamma t_i}{p+2\gamma t_j}\right)^{-\sigma}}{\partial p} = -\sigma \left(\frac{p+2\gamma t_i}{p+2\gamma t_j}\right)^{-\sigma-1} \left(\frac{1(p+2\gamma t_j) - 1(p+2\gamma t_i)}{(p+2\gamma t_j)^2}\right) > 0.$$

## Appendix B. Adding a nest for all other goods

To provide welfare results that account for substitution between restaurants and all other goods, I specify a nested-logit model with one nest for restaurants and one nest for all other goods. The linear utility specification is similar to that in the nested-logit model of Section 7, and generates choice probabilities which can easily be interpreted in terms of expenditure shares. An individual first solves the maximization problem within the nest for restaurants (exactly as in the logit model of Section 2) and within the nest for all other goods. Then he solves the aggregate utility maximization problem by choosing expenditure shares on restaurants and on all other goods. Denote the quantity of the restaurant good purchased in area $k$ by $R_k$, the quantity of all other goods by $G_k$, the price index for restaurant (from equation 6) by $P_{Rk}$ and the price index for all other goods by $P_{Gk}$. Solving the aggregate utility maximization problem leads to $R/G = (P_R/P_G)^{-\nu}$, where $\nu$ is the elasticity of substitution between restaurants and all other goods. Denote the price elasticity of demand for restaurant by $\epsilon_R$, such that $(\partial R/\partial P_R)(P_R/R) = \epsilon_R$. It is straightforward to show that $\epsilon = -\nu$, so that $\epsilon = -1$ corresponds to the limiting case $\nu = 1$.[84] The aggregate relative price index between area $k$ and $k'$ is:

$$A_{k,k'} = \frac{(P_{Rk'}^{1-\nu} + P_{Gk'}^{1-\nu})^{1/(1-\nu)}}{(P_{Rk}^{1-\nu} + P_{Gk}^{1-\nu})^{1/(1-\nu)}}.$$

As shown in Sato (1976) and Vartia (1976), one can express the relative price index above in terms of expenditure shares. For instance if $s_{Rk'}$ is the expenditure share on restaurants in area $k'$, then:

$$A_{k,k'} = \left(\frac{P_{Gk'}}{P_{Gk}}\right)^{w_{Gk'}} \left(\frac{P_{Rk'}}{P_{Rk}}\right)^{w_{Rk'}},$$

where:

$$w_{Rk'} = \frac{(s_{Rk'} - s_{Rk})/(ln(s_{Rk'}) - ln(s_{Rk}))}{(s_{Rk'} - s_{Rk})/(ln(s_{Rk'}) - ln(s_{Rk})) + (s_{Gk'} - s_{Gk})/(ln(s_{Gk'}) - ln(s_{Gk}))}.$$

[84] It is standard to assume $\nu > 1$, but the welfare estimates are not sensitive to using $\epsilon = -1.02$, the price elasticity of demand for food away from home suggested by Okrent and Alston (2010) meta-analysis, instead of $\epsilon = -1$.

I assume that the price index for all other goods is constant across areas, so that $P_{Gk} = P_{Gk'}$. As the price elasticity of demand for restaurants equals -1, the expenditure share on restaurants is constant, so that $s_{Rk} = s_{Rk'}$ for any areas $k$ and $k'$. If $s_{Rk'}$ is arbitrarily close to $s_{Rk}$, then we can find $w_{Rk'}$ as:

$$\lim_{s_{Rk} \to s_{Rk'}} w_{Rk'} = s_{Rk'}.$$

With data on expenditure shares, it is now possible to compute the aggregate relative price index and measure the average household's willingness to pay to prevent a 20% increase in the restaurant price index. The 2009 CEX suggests that food away from home, that I take as a proxy for restaurants, accounts for 5.3%, of total expenditures, so I set $s_{Rk'}$ = 0.053. The aggregate relative price index becomes $A_{k,k'} = \left(\frac{P_{Rk'}}{P_{Rk}}\right)^{w_{Rk'}} = \left(\frac{P_{Rk'}}{P_{Rk}}\right)^{w_{Rk'}} = 1.2^{0.053}$ = 1.0097. Average total household expenditures in the CEX 2009 is about \$49,000, so for an average household the willingness to pay to prevent a 20% increase in the restaurant price index is the absolute value of 49,000(1-1,0097), which equals \$475.

## Appendix C. The determinants of the probability of making a restaurant trip.

Regressions on the determinants of the probability of making a restaurant trip can tell us about the importance of substitution between restaurants and all other goods, and about the distribution of the gains from density across different types of households. The sample for these regressions consists of all individuals in the NHTS for the 15 states that I use in my analysis, i.e. of 250,309 individuals 5 years and older. Each individual in the sample is an observation, indexed by $j$. I create two dummy variables for use as dependent variables. The variable $D_j^{\text{ALL}}$ is equal to one if an individual took at least one trip to a restaurant, by any mode and starting from any origin, on his travel day. The variable $D_j^{\text{HOME}}$ is equal to one if an individual took at least one trip to a restaurant from home, by any mode, on his travel day.[85] I specify a linear probability model with either $D_j^{\text{ALL}}$ or $D_j^{\text{HOME}}$ as a dependent variable.[86] The estimating equation is:

$$D_j^{\text{ORIGIN}} = \alpha + \beta_1 loc_k + \beta_2 div_k + \beta_3 X_j + \epsilon_{jk} \qquad \text{ORIGIN} \in \{\text{ALL,HOME}\}, \tag{C1}$$

[85]Only 11% of the individuals who take at least one trip to a restaurant on their travel day take more than one trip, and my analysis ignores the possibility that individuals travel more than once to a restaurant on their travel day.

[86] I verify that the average marginal effects from a probit model generate almost the same coefficients.

where $loc_k$ represents measures of restaurant location in the area $k$ and individual $j$ lives in, $div_k$ is a control for restaurant diversity and $X_j$ is a vector of individual characteristics for individual $j$ such as age, gender, education, and income.

The regression results are in table 8. The dependent variable in the regressions of columns 1-3 is the dummy for any trip to a restaurant from all origins, and in columns 4-6 it is a dummy for any trip to a restaurant from home. To interpret the coefficients, note that the probability that an individual visits a restaurant on any given day (the mean of $D_j^{\text{ALL}}$) is 0.2, while the probability that an individual visits a restaurant from home (the mean of $D_j^{\text{HOME}}$) is 0.08. Also, recall that the measures of restaurant locations are computed starting from home. In column 1, the log of global density enters alone with a coefficient of 0.0056, so that doubling the number of restaurant within 45 minutes of travel increases the probability of at least on trip to a restaurant on any given day by about 0.0056 percentage point from a 0.2 percent basis i.e. by 2.8%. Similarly, the coefficient on global density of 0.004 in column 4 is equivalent to a 5% increase in the probability of at least one trip from home. The coefficient on global density becomes smaller or negative when adding controls for other measures of restaurant location (time to the closest restaurant, local density, skewness of the restaurant distribution) and specialization in columns 2-3 and 5-6 (in columns 3 and 6 I also add individual characteristics). Generally, the measures of restaurant location in the home area have only small effects on the probability of making a trip to a restaurant from any origin, but larger, more robust effects on the probability of a trip that originates directly from home. For instance, travel time to the closest restaurant has a small and positive effect on the probability of a trip from any origin, but it has the expected large negative effect on the probability of a trip from home. Doubling travel time to the closest restaurant decreases the probability that an individual takes at least one trip to a restaurant from home on any given day by 7.5%, a result robust to controling for individual characteristics. In section 6 individual characteristics were not important predictors of trip time, but they have much more explanatory power over the probability of a trip. Income is the main determinant of the probability of visiting a restaurant on any given day. Compared with individuals with median income ($45,000 to $55,000), individuals in the top income bracket (larger than $100,000) are 15% more likely to visit at least one restaurant on any given day. Individuals in the lowest income bracket (less than $20,000) are 35% less likely to visit at least one restaurant than individuals in households with median income. These results suggest that high income household have a significantly higher willingness to pay for restaurant density.

***Regression using the variety-adjusted price index as a regressor***

It is also useful to evaluate the impact of the variety-adjusted price index at home (which can be thought of as a measure of density with a microeconomic foundation) on the probability of a trip to a restaurant. The sample and estimating equation are exactly as in equation (C1), but now the only measures of restaurant location are 9 dummies for each decile of $P_k$, the restaurant price index from the logit model in the area $k$ individual $j$ lives in. I find that within the 9 lowest deciles of $P_k$ there is only negligible variation in the probability of a trip to a restaurant from any origin, and individuals in the $10^{th}$ decile are 10% less likely to visit at least one restaurant on any given day. However, changes in the price index have a larger impact on the probability of a trip from home. Individuals living in areas within the median decile of the price index (average price index equals to 10.0) are 11% less likely to travel to at least one restaurant from home than individuals living in areas within the lowest decile of the price index (average price index equals to 8.7). Individuals living in areas within the highest decile of the price index (average price index equals to 13.8) are 33% less likely to make at least one trip from home than individuals living in areas within the median decile. These numbers imply, for instance, that a 20% increase in the variety-adjusted restaurant price index leads to an approximately 17% decrease in the probability that an individual makes at least one trip to a restaurant from home on his travel day, and that most of this decrease is compensated by an increase in trips with an origin other than home.

## Appendix D. Nested model with supply-side

In this appendix, I derive $\frac{d_{kc}}{d_{kc'}} = \left(\frac{b_{kc}}{b_{kc'}}\right)^{\frac{\sigma-1}{\sigma-\mu}}$ an expression relating relative tastes for restaurant categories to the relative densities of restaurant in these categories. A model with a discrete number of restaurants and consumers who optimally choose their location would be analytically intractable, and cumbersome to estimate numerically. I therefore simplify the model by assuming that an infinity of restaurants and consumers are uniformly distributed on the line. This generates the intuitive, analytical expression above, which is based on theory but derived from slightly different assumptions that those underlying the model that I estimate.

Recall that the discrete version of the nested-CES model leads to exactly the same outcomes as a nested-logit model. So I solve a nested-CES model, but with subutilities (within-nest) defined

Table 8: Determinants of the probability of a trip

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dummy, trip to restaurant from: | ALL | ALL | ALL | HOME | HOME | HOME |
| log Global density | $0.0056^a$ | $0.0026^b$ | $-0.0017^a$ | $0.0013^c$ | $0.040^a$ | 0.0008 |
| | (0.0006) | (0.0011) | (0.0013) | (0.0005) | (0.0008) | (0.009) |
| log Skewness | | $-0.0060^a$ | $-0.0036^a$ | | $-0.0049^a$ | $-0.0047^a$ |
| | | (0.0009) | (0.0010) | | (0.006) | (0.007) |
| log Time to closest rest. | | $0.0050^b$ | 0.0011 | | $-0.0050^a$ | $-0.0059^a$ |
| | | (0.0017) | (0.019) | | (0.0012) | (0.0013) |
| log Local density | | $0.0056^a$ | $0.0039^b$ | | $0.0058^a$ | $0.0051^a$ |
| | | (0.0014) | (0.0017) | | (0.0010) | (0.0011) |
| log Specialization | | $-0.0751^a$ | $-0.0396^a$ | | $-0.0434^a$ | $-0.0273^a$ |
| | | (0.0107) | (0.0124) | | (0.0074) | (0.0273) |
| **Controls** | | | | | | |
| Individual characteristics | | | X | | | X |
| Observations | 250309 | 240029 | 188126 | 250309 | 240029 | 188126 |
| $R^2$ | 0.0003 | 0.0010 | 0.013 | 0.0004 | 0.0017 | 0.0087 |

*Notes:* OLS regressions with a constant in all columns. Robust standard errors in parentheses. $a$, $b$, $c$: significant at 1%, 5%, 10%. In columns 1-3 the dependent variables is a dummy for whether an individual traveled at least once to a restaurant on his NHTS travel day, and in columns 4-6 the dependant variable is a dummy for whether an individual traveled at least once to a restaurant, from home, on his NHTS travel day. On any given day, the probability of at least one restaurant trip from any origin is 0.2 and the probability of at least one restaurant trip from home is 0.08. Individual characteristics include 17 dummies for household income, 4 dummies for education, household size, 6 dummies for age, a dummy for gender, a dummy for black, and a dummy for worker's status.

over a continuous set of restaurants, and aggregate utility defined over a discrete number of categories. I consider a linear city of infinite length, with a continuous and uniform distribution of consumers, indexed by $j$. An exogenous parameter $d^s$ captures the density of consumers on the line.[87] Each consumer's value of travel time is $\gamma > 0$. Each restaurant meal is produced with fixed costs $F$ and marginal cost $mc$, and there there is free entry into the restaurant market. There are $C$ restaurant categories, indexed by $c$, and an infinity of restaurants in each category, uniformly located on the real line, and indexed by $i$.[88] The density of restaurant in each category, $d^r_c$, is endogenous.

The subutility of consumer $j$ within category $c$ is:

$$Q_{jc} = 2\int_0^{\infty} \left(q_{jc}(i)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \tag{D1}$$

[87]Strictly speaking $d^s$ cannot be a density or a parameter in a probability density function, because the uniform distribution is not defined on the real line. This does not pose any technical challenges however, and $d^s$ will capture density in an intuitive sense.

[88]I use the same index $i$ for each category, as there can be no possible confusion.

where $q_{jc}(i)$ is the number of trips (or meals) to restaurant $i$ in category $c$, $p_{jc}(i)$ is the price of that restaurant, including transport costs, and $\sigma > 2$ is the elasticity of substitution within categories. A factor 2 multiplies the integral because consumers have access to restaurants on both sides of their location (I make this precise when computing aggregate demand for each restaurant). The aggregate utility of a consumer $j$ is:

$$U_j = \sum_{c=1}^{C} \left( b_c^{1/\mu} Q_{jc}^{\frac{\mu-1}{\mu}} \right)^{\frac{\mu}{\mu-1}}, \tag{D2}$$

where $b_c$ is an exogenous taste parameter for category $c$, and $\mu > 2$ is an elasticity of substitution across categories. Maximizing the subutility in equation (D1) subject to the budget constraint $\int_0^\infty 2p_{jc}(i)q_{jc}(i)di = y_{jc}$, where $y_{jc}$ is consumer $j$'s expenditures on restaurants in category $c$, I obtain a standard CES demand function:

$$q_{jc}(i) = \frac{p_{jc}(i)^{-\sigma} y_{jc}}{P_{jc}^{1-\sigma}}, \tag{D3}$$

where:

$$P_{jc} = 2 \left( \int_0^\infty p_{jc}(i)^{1-\sigma} di \right)^{\frac{1}{1-\sigma}} \tag{D4}$$

is a CES price aggregator.

Maximizing the aggregate utility in equation (D2) subject to the budget constraint $\sum_{c=1}^{C} P_{jc}Q_{jc} = E$, where E is aggregate expenditures on restaurants. I obtain the demand function for the composite restaurant good:

$$Q_{jc} = \frac{b_c P_{jc}^{-\mu} E}{P_j^{1-\mu}}, \tag{D5}$$

where:

$$P_j = \sum_{c=1}^{C} \left( b_c P_{jc}^{1-\mu} \right)^{\frac{1}{1-\mu}} \tag{D6}$$

is a price aggregator.

A property of the price aggregator in equation (D4) is that $P_{jc}Q_{jc} = y_{jc}$, so that $\sum_{c=1}^{C} P_{jc}Q_{jc} = \sum_{c=1}^{C} y_{jc}$, and $\sum_{c=1}^{C} y_{jc} = E$, as it should. To find an expression for the demand of each consumer for each restaurant as a function only of restaurant prices and total expenditures, I first substitute $y_{jc} = P_{jc}Q_{jc}$ into equation (D3), and then I substitute the demand for the composite restaurant good in category $c$ by consumer $j$ (equation D5) into the demand for each restaurant $i$ (equation D3). I obtain the demand for restaurant $i'$ in category $c'$ by consumer $j$ as

$q_{jc'}(i') = \frac{Eb_{c'}p_{jc'}(i')^{-\sigma}P_{jc'}^{\sigma-\mu}}{P_j^{1-\mu}}$, which can also be written as:

$$q_{jc'}(i') = \frac{Eb_{c'}p_{jc'}(i')^{-\sigma}\left(2\int_0^\infty p_{jc'}(i)^{1-\sigma}di\right)^{\frac{\sigma-\mu}{1-\sigma}}}{\sum_{c=1}^C b_c\left(2\int_0^\infty p_{jc}(i)^{1-\sigma}di\right)^{\frac{1-\mu}{1-\sigma}}}. \tag{D7}$$

I now find the aggregate demand for each restaurant. Consider the price of restaurant $i$ in category $c$. A consumer living exactly at restaurant $i$'s location pays only the meal price $p_{ci}^{meal}$. I assign index $j$ = 0 to this consumer, so that $p_{0c}(i) = p_{ci}^{meal}$, and I define the price of restaurant $i$ for each consumer $j$ as $p_{jc}(i) = p_{ci}^{meal} + 2\gamma\frac{j}{d^s}$. One can think of consumer $j$ as living at distance $\frac{j}{d^s}$ from restaurant $i$, with $d^s$ capturing the density of individuals on the line.[89] Now consider restaurant prices again, but this time from the perspective of a consumer (say consumer $j'$). I assign an index $i$ = 0 to a restaurant located exactly where consumer $j'$ lives, so its price is $p_{j'c}(0) = p_{c0}^{meal}$. Generally the price of restaurant $i$ is $p_{j'c}(i) = p_{ci}^{meal} + 2\gamma\frac{i}{d_c^r}$, where $d_c^r$ is a measure of restaurant density.[90] I can rewrite the demand for restaurant $i'$ in category $c'$ by consumer $j$ as:

$$q_{jc'}(i') = \frac{Eb_{c'}\left(p_{c'i'}^{meal} + 2\gamma\frac{j}{d^s}\right)^{-\sigma}\left(2\int_0^\infty\left(p_{c'i}^{meal} + 2\gamma\frac{i}{d_{c'}^r}\right)^{1-\sigma}di\right)^{\frac{\sigma-\mu}{1-\sigma}}}{\sum_{c=1}^C b_c\left(2\int_0^\infty\left(p_{ci}^{meal} + 2\gamma\frac{i}{d_c^r}\right)^{1-\sigma}di\right)^{\frac{1-\mu}{1-\sigma}}}. \tag{D8}$$

To obtain the aggregate demand for restaurant $i'$ in category $c'$, I integrate equation (D8) over all consumers from location $j = 0$ to $j = \infty$, and multiply by 2 to account for the presence of consumers on both sides of restaurants $i'$. I obtain:

$$q_{c'i'} = \frac{Eb_{c'}\left(2\int_0^\infty\left(p_{c'i'}^{meal} + 2\gamma\frac{j}{d^s}\right)^{-\sigma}dj\right)\left(2\int_0^\infty\left(p_{c'i}^{meal} + 2\gamma\frac{i}{d_{c'}^r}\right)^{1-\sigma}di\right)^{\frac{\sigma-\mu}{1-\sigma}}}{\sum_{c=1}^C b_c\left(2\int_0^\infty\left(p_{ci}^{meal} + 2\gamma\frac{i}{d_c^r}\right)^{1-\sigma}di\right)^{\frac{1-\mu}{1-\sigma}}}. \tag{D9}$$

The profit-maximization problem of restaurant $i'$ in category $c'$ is:

$$\max_{p_{c'i'}^{meal}} q_{c'i'}(p_{c'i'}^{meal} - mc) - F. \tag{D10}$$

To simply the expression for $q_{c'i'}$, I solve the first integral in the numerator of equation (D9):

$$\int_0^\infty\left(p_{c'i'}^{meal} + 2\gamma\frac{j}{d^s}\right)^{-\sigma}dj = \frac{d^s(p_{c'i'}^{meal})^{1-\sigma}}{2\gamma(\sigma-1)},$$

[89] I could define an index for consumers on the right of restaurant $i$ and an index for consumers on its left, but the results are clearer keeping only one index $j$.

[90] Note that there are restaurants on both sides of consumer $j'$, which is why I multiplied the CES subutility by 2 in equation (D1).

and I replace everything that does not depend on the choice variable $p^{meal}_{c'i'}$ by a constant X, so that $q_{c'i'} = X(p^{meal}_{c'i'})^{1-\sigma}$ and the profit maximization problem becomes:

$$\max_{p^{meal}_{c'i'}} X(p^{meal}_{c'i'})^{1-\sigma}(p^{meal}_{c'i'} - mc) - F,$$

which has a maximum[91] at:

$$p^{meal}_{c'i'} = \frac{(1-\sigma)mc}{2-\sigma}. \tag{D11}$$

I now use a free-entry condition to derive an expression for the relationship between exogenous tastes $b_c$ and restaurant density $d^r_c$ in each category. From equation (D11), meal prices are constant across categories, so in a zero-profit equilibrium it must be that $q_{c''i''} = q_{c'i'}$ for any categories $c$ and $c'$ (profits do not vary within category by construction). Solving both integrals in the numerator of equation (D9), aggregate demand for restaurant $i'$ in category $c'$ becomes:

$$q_{c'i'} = \frac{Eb_{c'}\left(\frac{2d^s(p^{meal}_{c'i'})^{(1-\sigma)}}{2\gamma(\sigma-1)}\right)\left(\frac{2d^r_{c'}(p^{meal}_{c'i'})^{(2-\sigma)}}{2\gamma(\sigma-2)}\right)^{\frac{\sigma-\mu}{1-\sigma}}}{\sum_{c=1}^{C} b_c\left(2\int_0^\infty\left(p^{meal}_{ci} + 2\gamma\frac{i}{d^r_c}\right)^{1-\sigma} di\right)^{\frac{1-\mu}{1-\sigma}}},$$

and I can write $q_{c'i'} = q_{c''i''}$ as:

$$b_{c'}\left(\frac{2d^r_{c'}p^{meal}_{c'}(i')^{(2-\sigma)}}{2\gamma(\sigma-2)}\right)^{\frac{\sigma-\mu}{1-\sigma}} = b_{c''}\left(\frac{2d^r_{c''}p^{meal}_{c''}(i'')^{(2-\sigma)}}{2\gamma(\sigma-2)}\right)^{\frac{\sigma-\mu}{1-\sigma}}$$

which further reduces to:

$$\frac{d^r_{c'}}{d^r_{c''}} = \left(\frac{b_{c'}}{b_{c''}}\right)^{\frac{\sigma-1}{\sigma-\mu}},$$

the condition that I use to estimate the nested-logit model in section 7. Note that in the main text the $r$ superscript is omitted, and there is an additional area specific index $k$.[92]

[91] To see that $p^{meal}_{c'i'} = \frac{(1-\sigma)mc}{2-\sigma}$ is a maximum for $p^{meal}_{c'i'} \geq 0$, note that term $X(p^{meal}_{c'i'})^{1-\sigma}(p^{meal}_{c'i'} - mc)$ tends to negative infinity as $p^{meal}_{c'i'} \to 0$, tends to 0 as $p^{meal}_{c'i'} \to \infty$, and has a positive value for all $\sigma > 2$ at its only stationary point $p^{meal}_{c'i'} = \frac{(1-\sigma)mc}{2-\sigma}$.

[92] Setting profits to zero, it is possible to obtain an expression for the density of restaurants in each category:

$$d^r_{c'} = \frac{Ed_s(\sigma-2)}{F(\sigma-1)^2\left(1+\sum_{c=1}^{C}\left(\frac{b_{c'}}{b_c}\right)^{\frac{1-\sigma}{\sigma-\mu}}\right)}.$$

If there is only one category, this expression reduces to:

$$d^r = \frac{Ed_s(\sigma-2)}{F(\sigma-1)^2},$$

and it is straightforward to show that $\frac{\partial d^r}{\partial\sigma} < 0$ for $\sigma > 2$ . This relationship between restaurant density and the elasticity of substitution is a potential source of endogeneity bias in my estimates of $\sigma$.

## Appendix E. List of restaurant categories

Table 9 lists the restaurant categories defining the 85 nests in the nested-logit model, and their percentage share in the restaurant sample. The definition of categories is that from Yelp.com at the time of data collection. I use regular expressions to match each Google Places restaurant category to a closely related Yelp category. Almost 17% of restaurants on Google Places have category 'undefined', almost always because there was no information on the category of these restaurants on Google. These places are generally independent restaurants serving standard fares, and they are usually not on Yelp. Unsurprisingly, the definition of very small categories is almost irrelevant, so consolidating the 85 categories into 30 (for instance by merging Halal, Persian\Iranian, Middle Eastern, Moroccan and Turkish) or even 8 categories has little impact on the estimation results.

Table 9: Restaurant categories and percentage share

| Category | Percentage Share (%) | Category | Percentage Share (%) |
|---|---|---|---|
| Undefined | 16.7 | Cajun\Creole | .02 |
| Pizza | 13.3 | Ethiopian | 0.02 |
| Mexican | 9.3 | African | 0.02 |
| American | 9.1 | Filipino | 0.01 |
| Burger | 7.5 | Persian\Iranian | 0.01 |
| Chinese | 6.3 | Turkish | 0.01 |
| Sandwich | 5.2 | Peruvian | 0.01 |
| Cafe | 4.1 | Soup | 0.01 |
| Deli | 3.6 | British | 0.01 |
| Pub | 3.6 | Fondue | 0.01 |
| Seafood | 2.8 | Tapas | 0.009 |
| Italian | 2.6 | Portuguese | 0.008 |
| Chicken | 2.4 | Hawaiian | 0.008 |
| Barbecue | 2.3 | Mongolian | 0.008 |
| Steak | 1.7 | Southern | 0.008 |
| Japanese | 1.7 | Modern European | 0.006 |
| Thai | 1.0 | Halal | 0.006 |
| Diner | 0.8 | Russian | 0.006 |
| Fast | 0.7 | Pakistani | 0.006 |
| French | 0.6 | Polish | 0.005 |
| Indian | 0.6 | Moroccan | 0.005 |
| Greek | 0.5 | Afghan | 0.004 |
| Breakfast | 0.4 | Argentine | 0.003 |
| Vietnamese | 0.4 | Live\Raw | 0.003 |
| Brewery | 0.3 | Gastropubs | 0.003 |
| Sushi | 0.3 | Belgian | 0.003 |
| Hot Dog | 0.2 | Tex-Mex | 0.002 |
| Buffet | 0.2 | Malaysian | 0.002 |
| Asian Fusion | 0.2 | Cheesesteak | 0.002 |
| Korean | 0.2 | Gluten-Free | 0.002 |
| Mediterranean | 0.1 | Taiwanese | 0.002 |
| Irish | 0.1 | Crepe | 0.001 |
| Spanish | 0.1 | Burmese | 0.001 |
| Vegetarian | 0.1 | Carribean | 0.001 |
| Soul Food | 0.1 | Indonesian | 0.0009 |
| Middle Eastern | 0.09 | Hungarian | 0.0009 |
| Latin American | 0.06 | Cambodian | 0.0009 |
| German | 0.05 | Basque | 0.0007 |
| Cuban | 0.05 | Himalayan\Nepalese | 0.0005 |
| Vegan | 0.03 | Scandinavian | 0.0002 |
| Fish & Chips | 0.02 | Ukrainian | 0.0002 |
| Kosher | 0.02 | Singaporean | 0.0002 |
| Brazilian | 0.02 | | |

*Notes:* To accommodate differences in the Yelp and Google Places terminology, I make the following modifications to the original set of Yelp categories: I merge the 'New American' and 'Traditional American' categories in a single 'American' category, I drop the category 'Food Stand', I divide the Brewery\Grill & Pub category into two categories: 'Brewery' and 'Pub', and I merge the 'Tapas Bar' and 'Small Plates Tapas' categories into a single 'Tapas' category.