Quality of Consumption: A Revealed Preference Approach to Local Amenity Valuation

Benjamin Bridgman, Abe Dunn and Mahsa Gholizadeh*

March 26, 2026

Abstract

Traditional revealed-preference measures of quality-of-life from a particular location are inferred by assuming individuals choose locations optimally in equilibrium and inferring the quality-of-life by comparing cost-of-living in an area relative to wages. Similarly, the quality of consumption in one location over another may also be revealed from an individual's decisions to visit and consume in one location over another. We present a simple model of consumption preferences across locations, which we use to estimate "consumption amenities" of different locations across the U.S. We use novel data of credit card transactions to measure these amenities using discrete choice demand models, which show varied consumption patterns across industries. We show that the consumption amenities are highly correlated with housing prices and other measures of amenities. However, we also identify high quality of consumption areas that do poorly in the traditional methods. We also show that social frictions, the tendency for similar demographics to consume together, are the same magnitude as distance frictions.

*We thank the seminar participants at the NBER Disaggregated National Accounts: Measurement and Application conference. The views expressed in this paper are solely those of the authors and not necessarily those of the U.S. Bureau of Economic Analysis or the U.S. Department of Commerce.

1 Introduction

The quality of living (QOL) of a location is a foundational question in urban and regional economics. Buying or renting housing in a location gives access to a bundle of attributes, including the local labor market, natural amenities like weather, and produced amenities like retail outlets or public goods.

While data on housing and labor market returns (wages) are abundant, amenities are harder to value. Many do not have a direct price to compare with the attributes that do have a money price. A common method around this problem is to back out amenity value as a residual: the value of housing cost that is not explained by wages adjusted by indicators of amenities. There is a long literature that constructs quality of life estimates for cities based on these equilibrium assumptions Roback (1982), Blomquist, Berger and Hoehn (1988), and Albouy (2019). These methods have challenges. Housing prices are influenced by supply restrictions that are not related to amenities. Amenities come as bundles, which introduces imperfect identification of the contribution of each amenity. Measuring the contribution of amenities is becoming more important since they are becoming stronger influence in location decisions (Glaeser et al. 2001, Rappoport 2008, Albouy 2019, Couture and Handbury 2020).

This paper estimates an important amenity: quality of consumption (QOC) in the local area. People can travel to experience and consume in non-home areas. People consume a lot outside their home county even for "non-traded" services (Dunn and Gholizadeh 2024). Batch et al. (2025) find that 32 percent of expenditure on services that require physical presence is outside consumers' home county. This indicates that people do not see outlets across locations as equivalent: they expend time and money to pass competing outlets when shopping.

We estimate local area QOC using revealed preferences on credit card data. High quality consumption locations are those that have higher sales relative to what would be expected. Conditional on choosing a home location, individuals choose to consume in different locations, which will depend on the travel time, prices of goods and services in these locations, and also sociodemographic factors that may influence travel patterns. The value of consumption in a location is taken from the common component of demand across individuals, after netting out

other factors, such as distance, price, and sociodemographic factors shaping consumer choices. This method in valuing amenities does not rely on housing prices so provides independent information about their value.

To estimate these quality measures, we use credit card transaction data from FirstData, one of the largest credit card intermediaries in the country. The data was managed and processed by Palantir to provide aggregated and anonymized information of spending flows across geographies by three digit NAICS industry codes. These data are combined with Census business establishment data along with other sources used in the construction of regional GDP statistics to build representative estimates of total consumption flows across areas over the time period from 2012 to 2017.

We examine the geographic distribution of expenditures across 17 three digit NAICs categories that account for a total of about 70 percent of consumer spending nationally. The geographic unit of analysis in this paper is the county, which we aggregate into metropolitan and non-urban state areas. This aggregation allows us to compare our findings with the urban economics literature, which focuses on metropolitan areas as the unit of observation.

Our measure of consumption quality gives reasonable rankings. The highest ranking QOC cities generally match to other rankings of appealing locations. The highest ranked locations are Honolulu and cities in the Sun Belt.

The cost of housing reflects this attractiveness, as there is a strong correlation between housing cost and QOC. However, we do not simply recover housing costs. The deviations show the value of an independent measure. We find several low housing cost, high wage, and high QOC locations, including fast growing Texas cities like Dallas and Houston. Methods using residuals of wages and housing prices give them low rankings. They need a low amenity value to equalize locational value.

We also document how geographic and social frictions shape consumption activity. Our findings consistent with both distance and social frictions affecting economic activity (e.g., race, education, and political affiliation). We find that social frictions are of a similar magnitude as distance. Our paper is the first to document these frictions identified by Davis et al. (2019) using consumption expenditure data.

2 Related Literature

The geography of consumption is fundamental part of measuring and understanding regional economies and constructing regional accounts. The total consumption of individuals that reside in an area equals total consumption produced in an area plus net exports (i.e., the total amount sold to individuals outside of the area, plus total amount purchased outside of the area by individuals that reside in an area), but the data to look at this relationship has not been available until relatively recently.[1] This paper contributes to a growing literature using alternative data to measure the geography of consumption, including Batch et al. (2025) and Blanchard et al. (2025).

This paper is related to the literature on local economic shocks and spillovers, as consumption may be one avenue for geographic spillovers to occur (e.g., Nakamura and Steinsson (2014), Aurerback and Gorodnichenko (2012), Aurerback and Gorodnichenko (2019), and Dunn and Gholizadeh (2025).). In addition to distance being a factor determining consumption across areas, we find evidence that socioeconomic factors may play a role in determining where individuals consume. This finding relates to work by Davis, Dingel, Monras and Morales (2019) who use Yelp data to show evidence of consumption segregation for the restaurant sector. Our paper supports their finding, and shows that the consumption segregation exists across a number of industries based on actual transaction data, rather then just restaurant reviews. Of course, there are numerous factors in addition to consumption linking geographies. Monte, Redding, and Rossi-Hansberg (2018) consider linkages in goods and factor markets across areas where they examine labor demand shocks on employment elasticities using a general equilibrium framework. Across-market consumption is an additional factor linking markets that has not previously been considered due to data limitations.

This paper contributes to a growing literature using revealed preference to estimate the value of locations, including Mast and Barca (2026) and Bell, et al. (2026). Our paper applies these concepts to local consumption.

The method used to infer location quality are quite similar to those applied in the discrete

[1] Dolfen, Einav, Klenow, Klopack, Levin, Levin and Best (2019) using Visa data to look at the gains of E-commerce and growth in E-commerce over time. In contrast, this paper focuses explicitly on brick and mortar stores in a single cross section.

choice and industrial organization literature, McFadden (1974, 1978) and Berry (1994) , where the unobserved quality of differentiated products are revealed through the choices of individuals in the market. Many papers in the industrial organization literature find that geographic differentiation of establishments is an important dimension of differentiation and is an important factor determining competition among establishments (e.g., Houde (2012) - gasoline stations, Davis (2006) - movie theaters). This has been especially important for the health care sector, where geographic proximity to patients is a key element of any demand study (see Gaynor, Ho, and Town (2015) for a review). Unlike these papers, the "products" in this paper are three digit NAICS aggregations of establishments in a particular county, so the geographic units studied are much more coarse, and unlike these papers we are considering multiple industries, rather than just a single industry. A related literature looks at the decision of households to select a single residence locations among many (e.g., McFadden (1978) and Bayer, McMillan, Murphy and Timmins (2019)), considering the amenities and values across locations as an important factor. Unlike this branch of the literature, our paper focuses on the consumer's decision to consume in a location conditional on consumers already selecting their residence.

3 Basic Theory and Empirical Model

Previous work pioneered by Rosen (1979) and Roback (1982) used revealed preference to rank the value of different locations assumes that the location of individuals is based on equilibrium conditions determined by the wage differences across locations, cost-of-living, and the quality-of-life (QOL) across areas. Roughly, the marginal willingness-to-pay for the QOL in area j may be equated to a function of the cost-of-living and income. Specifically applied to U.S. data, Albouy (2012) estimates this relationship to be:

$$\widehat{QOL}_j = 0.33 \cdot \widehat{p}_j - 0.51 \cdot \widehat{w}_j$$

where $\widehat{QOL}_j$ is the estimate of the quality-of-life, $\widehat{p}_j$ is a normalized measure of the cost-of-living and $\widehat{w}_j$ is a normalized measure of incomes across areas. These values take into account federal

housing, non-housing costs, and non-labor income.

In this paper, we take the location of where individuals decide to live as given, and examine the determinants of consumption across locations to reveal information about the quality-of-consumption (QOC) across areas. The main idea is that even after individuals choose to locate in a particular location, they can travel to other location to experience many of the amenities across geographies, even if those amenities do not exist in their local area.

To determine the value of consumption across locations, we follow the discrete choice literature and treat consumption in different locations as differentiated products. Intuitively, one can view the method as allowing all consumers in the United States to choose the area in which they like to consume, based on where we observe them actually consuming and the associated travel cost for reaching that location. However, the choice will be highly skewed based on the location of individual consumers, so we use the model to remove these home-location effects and isolate the quality of consumption, regardless of the consumer's location.[2]

While we will expand this analysis to studying a variety of industries, for simplify the introduction of key concepts we will begin by focusing on just one industry, where we assume that consumers and producers have already selected locations. Based on these fixed locations, let the indirect utility of consumer i that resides in location l and consumes in location j be denoted $U_{ilj}(t(l,j), p_j, Q_j))$ where $t(l,j)$ is the transportation cost of the consumer from traveling from location l to location j, p_j is the price level of consumption in that location, the value x_k are other factors that may affect consumption across location, and Q_j is the quality of consumption (QOC). The quality of consumption Q_j will be estimated as a fixed effect, and $\xi_{l,j}$ is the residual quality term measured as the deviation from the fixed effect Q_j. On any particular day of the year, the consumer may choose to consume in location, j, if it provides them with higher utility than all other locations $k \in L$ where L denotes the set of all possible locations. That is, they consume in location l if $U_{ilj} > U_{ilk}$ for all $k \in L$. We will focus on a particular functional form of this indirect utility function, such that,

[2] A similar intuition is taken by Bayer, McMillan, Murphy and Timmins (2016) who use the discrete choice of residential location to infer the demand for houses and neighborhoods, including the unobserved quality of neighborhoods.

$$U_{ilj}(t(l,j),p_j,Q_j)) = t(l,j) + \beta_i \cdot x_j + \alpha \cdot p_j + Q_j + \xi_{l,j} + \epsilon_{ilj}$$

where ϵ_{ilj} takes on the type 1 extreme value distribution and we will normalize the value of this outside good to zero.

In this case, the probability that a consumer located in location l consumes a product in location j will take on the conditional-logit form, so the probability of consumer i residing in location l and consuming in location j may be expressed as:

$$\Pr(U_{ilj} > U_{ilk} \forall k \in L) = \frac{\exp(t(l,j) + \beta_i \cdot x_j + \alpha \cdot p_j + Q_j + \xi_{l,j})}{1 + \sum_{\forall k \in L} \exp(t(l,k) + \beta_i \cdot x_k + \alpha \cdot p_k + Q_k + \xi_{l,k})}$$

Assuming that all individuals in location l are identical, except for the idiosyncratic component ϵ_{ilj}, then we can express this probability in terms of market shares. That is, let the share of consumption of a consumer residing in location l and consuming in location j be denoted S_{lj}, then the expected probability of consumption for the typical individual in location l is equal to the share.

$$S_{lj} = \frac{\exp(t(l,j) + \beta_i \cdot x_{jl} + \alpha \cdot p_j + Q_j + \xi_{l,j})}{1 + \sum_{\forall k \in L} \exp(t(l,k) + \beta_i \cdot x_k + \alpha \cdot p_k + Q_k + \xi_{l,k})}$$

The empirical strategy will be to estimate the above model based on the observed flows of consumption across geographic locations, with the goal of estimating the value of Q_j. To do this, we will need to extract differences in consumption due to other factors, such as the price of the goods in an area, as well as transportation costs across areas.

Let the S_{l0} be the share of purchases for the outside good, which has utility normalized to zero.

$$S_{l0} = \frac{1}{1 + \sum_{\forall k \in L} \exp(t(l,k) + \beta_i \cdot x_k + \alpha \cdot p_k + Q_k + \xi_{l,k})}$$

The model can be made linear by taking the difference in the log of the shares following Berry (1994):

$$\ln(S_{lj}) - \ln(S_{l0}) = t(l,j) + \beta_i \cdot x_j + \alpha \cdot p_j + Q_j + \xi_{l,k}$$

After controlling for prices in the area (p_j), transportation costs ($t(l,j)$), and other observable socio-demographic factors ($\beta_i \cdot x_k$), the determinants of consumption in a location will depend on the quality of that consumption, Q_j, measured as a residual. This will, in part, be determined by the quality of the goods and services offered in that location and also the local amenities in an area that make consumption generally attractive. We refer to the combination of the quality of the goods, services and local amenities that enter Q_j as consumption amenities or, QOC. Following the insights of Horatscu and Joo (2016), a very similar functional form and insights may be motivated through a CES demand model, which is a common functional form applied in the trade literature. We will explore the robustness of our work to alternative functional forms and demand models with similar intuition driving these alternative models.

3.1 Multiple Industries

Unlike other studies that focus on a single industry, this paper has information on many 3 digit NAICS sectors. This paper focuses on the demand for consumption jointly across multiple sectors. The theoretical reason for considering the sectors jointly is that there may be agglomerations in consumption, since for a single trip consumers may save in transportation costs by purchasing across multiple industries. In other words, consumer's travel to a location they pay the same fixed transport cost no matter how many goods and services are consumed for a single trip. For example, a consumer may go to a grocery store, purchase gasoline, and go out to a restaurant, all in the same area.

Let each NAICS code industry be denoted by n, then the indirect utility function of the consumer is then $U_{iljn}(t(l,j), p_j, Q_j)) = t_n(l,j) + \beta_n \cdot x_{jl,n} + \alpha_n \cdot p_{j,n} + Q_j + \xi_{l,j,n} + \epsilon_{ilj}$. The mean utility is then defined by the following linear relationship:

$$\ln(S_{l,n,j}) - \ln(S_{l,n,0}) = t_n(l,j) + \beta_n \cdot x_{j,n} + \alpha_n \cdot p_{j,n} + Q_j + \xi_{l,j,n}$$

In addition to the theoretical justification, there is also an empirical justification for focusing on the average quality, Q_j, as measurement error from any one industry is less likely to affect the overall average. The basic idea is that factors that may create measurement error for one sector or location may be uncorrelated with the measurement error for other sectors.

4 Data

The main data source used in this paper is from Fiserv, a global card transaction intermediary, which processes $2 trillion dollars of card transactions for establishments around the world, including credit, debit, and prepaid gift cards that includes all card merchants (e.g., Visa, MasterCard, Discover and others).[3] First Data works with a company, Palantir, which is a technology company that specializes in the management and analysis of big data. Palantir has aggregated and anonymized transaction data in a way that provides detailed and meaningful economic information, while still protecting both merchant and individual anonymization. While some firms are excluded, there is still a large subset of firms that are available for analysis, including hundreds of thousands of merchants and billions of transactions that span all states in the U.S. including the District of Columbia.

The specific data we are analyzing here includes aggregate county-level information on establishments by three-digit North American Industry Classification System (NAICS) industry. For every county-NAICS combination, the data contains an estimate of the share of revenues that go to consumers residing across counties within the U.S. and also information about cards from foreign countries. For instance, the share of hotel revenues that go to Clark County, Nevada (i.e. Las Vegas) for the the NAICS category of Accommodations. For counties within the U.S., the card-holder's home location is estimated based on the transaction history of the consumer using information on all transactions across all industries available in the First Data database for the specific card.[4]

The data obtained from First Data contains only revenue shares of transactions between

[3] Other electronic card transactions are also included, such as Electronic Benefit Transfer (EBT).

[4] As a check on the home location algorithm, we are also provided with county-level data based on a subset of cards that have a known home-location zip-code.

county-to-county pairs between establishments and consumers across areas. To make these transaction shares representative of the total transactions in the area, we use a variety of data sources to produce revenue estimates for select industries from 2012 to 2017. Following the methodology of BEA for the construction of the regional economic accounts, we use a variety of data sources to estimate spending (i.e., nominal output) across industries at the county level. One of the key data sources is the Geographic Area Series of the 2012 Economic Census (EC) that is used which contains information on revenues and establishment counts by NAICS industry and county-level geographies. County-level receipts sometimes contain suppressions at the three digit industry level, which are imputed using other sources of data to create estimates for all counties in the United States in 2012. For the years 2012 forward, we use a combination of the National Economics Accounts data and detailed county-level information from the Quarterly Survey of Employment and Wages (QCEW) to estimate county-level growth rates by NAICS. The growth rates are used to project the 2012 economic census forward for all counties for the period from 2012 to 2017.[5]

Information on the transaction flows across geographies are suppressed in some cases, due to the processing of card transactions to protect the anonymity of merchants and consumers. This is especially common in area where revenues for the industry for a particular county are small. To generate a network of all transactions across the U.S. we apply a series of flexible models based on observable transactions in the database across areas, to generate estimates of transaction flows in nearly all areas where final goods are sold for our 17 select industries. Typically the spending information for those industries where transactions are observed in a county, combined with information on distances traveled for consumption, are used to impute these remaining spending flows. For instance, if we are missing accommodation flows in an area, but we observe flows of restaurant services, we can use information on the restaurant service flows between areas, combined with information on how far individuals typically travel to purchase accommodation services, to impute the flows for accommodation services between areas. The amount of imputation necessary varies by industry, but we have relatively good coverage for the

[5]A similar method is applied in estimating the regional economic accounts at BEA for some sectors, which is a methodology that has been developed, tested and refined over a number of years.

17 select 3 digit industries that we focus on in this study. The details of dollars observed by by industry and corresponding amounts with observed consumption flows and those imputed are shown in Table 1. Overall we observe consumption flows for 85.8 percent of final goods sold, we impute dollars transacted for 13.6 percent of final goods sold, and we are missing this information for less than 1 percent of final goods sold. For miscellaneous retail stores we observe consumption flows for 97 percent of dollars in final goods transacted, but we observe just 64 percent for health and personal care stores. Additional details describing the imputation process are included in the appendix.

Table 1: Total Receipts in Millions and the Associated Flows that are Observed, Imputed, and Unknown

	Observed	Imputed	Unknown
Accommodation	182.8	31.5	1.4
Ambulatory Health Care Services	889.9	38.5	4.5
Amusement, Gambling, and Recreation Industries	100.8	15.1	1.1
Building Material and Garden Equipment and Supplies Dealers	209.7	100.9	5.5
Clothing and Clothing Accessories Stores	240.5	11.1	1.7
Electronics and Appliance Stores	81.2	31.2	2.8
Food Services and Drinking Places	604.7	10.9	0.7
Food and Beverage Stores	599.7	83.1	1.6
Furniture and Home Furnishings Stores	84.8	19.7	1.2
Gasoline Stations	525.9	106.0	2.5
General Merchandise Stores	448.8	224.2	9.1
Health and Personal Care Stores	189.4	103.6	11.6
Miscellaneous Store Retailers	100.5	5.0	0.4
Performing Arts, Spectator Sports, and Related Industries	61.7	33.5	3.0
Personal and Laundry Services	95.4	5.3	0.7
Repair and Maintenance	151.7	16.4	1.8
Sporting Goods, Hobby, Book, and Music Stores	73.0	16.5	1.0

After generating estimates of the complete economic Census and merchant revenue flows across counties for a set of 3-digit NAICS industries, we combine these estimates together to form complete estimates of consumption flows across county-to-county pairs for the select 3 digit industries. Specifically, we multiply the share of NAICS revenue corresponding to a county-to-county pair times the estimated revenue for that NAICS to generate an estimate of the amount of consumer spending for consumers in that area. For example, for those consumers living in

Montgomery County, Maryland, we have an estimate of the dollars they spend on restaurants for every county in the United States.

We do not observe exact information on price, although we do observe an estimate of the geometric mean transaction value for every NAICS county-to-county pair in the data, which we use as a proxy for price. We use this information to estimate an average price per transaction for each merchant location and NAICS in our sample. Again, for areas where price information is suppressed, we impute missing prices using transaction amounts from other industries observed in that location. Additional details of the imputation process are described in the appendix. The measurement of price raises some concern over both measurement error and endogeneity of the estimated price, which we discuss in the empirical model section below.

We combine this information with several additional data sources. We use estimates of latitude and longitude for every counties population centroid from the Census Bureau. We also use both population and income information from the Bureau of Economic Analysis (BEA), as well as estimates of county demographics for all counties in the country from the Area Resource File.

5 Descriptive Statistics

The share of consumption at the home location relative to other locations varies greatly by industry and by location. Table 2 shows some basic descriptive statistics of the distribution of spending in a consumer's own home county. The share of spending occurring outside of a home county varies greatly by both industry and by county. Most counties in the United States are relatively rural, leaving individuals in the typical rural county to drive to more distant urban areas to purchase goods and services.

Table 3 shows the distribution of spending by distance, where the estimates in this table are weighted by dollars of final goods sold. When weighted by spending, we see that there is a much larger average share of spending occurring in the home county. At the same time, we still see a relatively large share of spending occurring outside of the home county. For those spending outside of their home county, most of the spending is occurring in a relatively short distance from

Table 2: Share of Spending from Merchant's Home County

	Median	Mean	25th	75th
Accommodation	0.007	0.041	0.002	0.035
Ambulatory Health Care Services	0.663	0.524	0.057	0.767
Amusement, Gambling, and Recreation Industries	0.045	0.257	0.008	0.527
Building Material and Garden Equipment and Supplies Dealers	0.036	0.301	0.010	0.645
Clothing and Clothing Accessories Stores	0.057	0.222	0.007	0.451
Electronics and Appliance Stores	0.031	0.174	0.010	0.264
Food Services and Drinking Places	0.555	0.501	0.348	0.675
Food and Beverage Stores	0.570	0.442	0.006	0.764
Furniture and Home Furnishings Stores	0.036	0.210	0.013	0.462
Gasoline Stations	0.480	0.423	0.158	0.666
General Merchandise Stores	0.011	0.246	0.005	0.571
Health and Personal Care Stores	0.036	0.260	0.013	0.611
Miscellaneous Store Retailers	0.363	0.338	0.020	0.557
Performing Arts, Spectator Sports, and Related Industries	0.010	0.060	0.004	0.032
Personal and Laundry Services	0.634	0.474	0.030	0.764
Repair and Maintenance	0.644	0.491	0.032	0.763
Sporting Goods, Hobby, Book, and Music Stores	0.022	0.180	0.006	0.362

their home location. The one industry that is the exception is accommodations, where most of the spending occurs in counties that are 100 miles or more from the home county.

6 Spatial Input-Output Linking Consumption and Final Goods Production

The data constructed for this study is unique as it connects the location of consumers with the location of final goods sold. Before turning to our specific application, we first want to demonstrate that these flows are a fundamental part of economic growth. We do this by establishing a basic relationship between total consumption in a county and final goods sold in that county minus net exports of consumption in that county (ignoring tax associated transfers). Simply stated, as an economy grows, the growth in consumption must occur somewhere, either through increased sales in the county or through an increase in sales outside of the county. The specific accounting relationship for a consumer in location i:

Table 3: Share of Spending in Merchant's Home County and Distance from Home County (Weighted by Spending)

	Home Share	Under 100	100 to 500	500+
Accommodation	0.117	0.125	0.326	0.431
Ambulatory Health Care Services	0.675	0.222	0.044	0.060
Amusement, Gambling, and Recreation Indu	0.541	0.202	0.107	0.150
Building Material and Garden Equipment a	0.521	0.191	0.135	0.153
Clothing and Clothing Accessories Stores	0.544	0.254	0.086	0.116
Electronics and Appliance Stores	0.487	0.184	0.127	0.202
Food Services and Drinking Places	0.666	0.206	0.062	0.066
Food and Beverage Stores	0.751	0.106	0.062	0.081
Furniture and Home Furnishings Stores	0.500	0.240	0.110	0.150
Gasoline Stations	0.616	0.173	0.112	0.099
General Merchandise Stores	0.513	0.148	0.155	0.184
Health and Personal Care Stores	0.476	0.179	0.149	0.197
Miscellaneous Store Retailers	0.593	0.196	0.084	0.128
Performing Arts, Spectator Sports, and R	0.274	0.199	0.190	0.337
Personal and Laundry Services	0.706	0.177	0.050	0.068
Repair and Maintenance	0.676	0.192	0.064	0.068
Sporting Goods, Hobby, Book, and Music S	0.491	0.213	0.128	0.168

$$\text{Consumption}_i = \text{Final Products Sold}_i - \text{Exports of Consumption}_i + \text{Imports of Consumption}_i$$

The exports and imports can be measured using a combination of flow data and final products sold described in Equation 1.

$$\begin{aligned}\text{Consumption}_i = \text{Final Products Sold}_i &- \sum_{j \neq i} \text{Share Merchant Export}_{j \to i} \cdot \text{Final Products Sold}_j \\ &+ \sum_{j \neq i} \text{Share Merchant Export}_{i \to j} \cdot \text{Final Products Sold}_i \end{aligned} \tag{1}$$

In words, the net exports in county i is measured by subtracting off consumption from consumers purchasing final products in county i that reside outside county i, but adding back in final products sold in counties outside of county i that are sold to individuals residing in county i.

This relationship cannot be observed using traditional data sources, as there is no information on exports or imports of consumption, but using credit card data we estimate this consumption-flow relationship.

To summarize the importance of these flows, we've calculated share of net exports to final products sold for our 17 select industries calculated as:

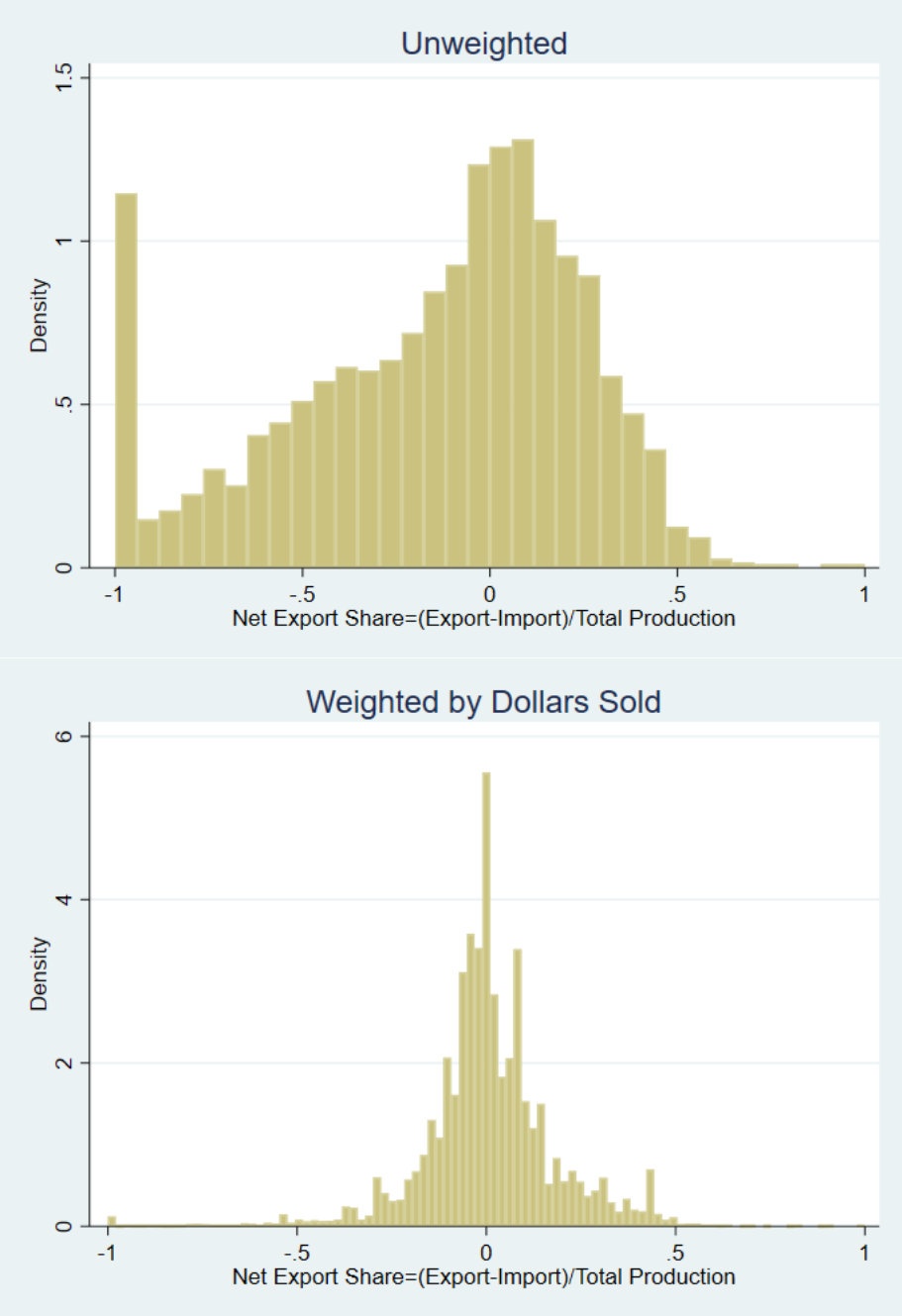
$$\text{Share Net Export}_i = \frac{\sum_{j \neq i} \text{Share Merchant Export}_{j \to i} \cdot \text{Final Products Sold}_j - \sum_{j \neq i} \text{Share Merchant Export}_{i \to j} \cdot \text{Final Products Sold}_i}{\text{Final Products Sold}_i} \tag{2}$$

where a negative share captures consumers in county i purchasing outside of their home county and positive share indicates more consumers entering the county to consume relative to leaving. A histogram of this share is plotted in Figure 1 where the data has been winsorized at -1. We see a large mass at -1 capturing many relatively small counties where most final goods purchased are outside of the immediate county, where it is likely that there is relatively little retail presence. In addition to the mass at -1, what is most striking is the variation in net exports across counties, which indicates that ignoring these net exports may be especially problematic for the study of consumption.

To empirically test the importance of this relationship, we substitute associated data elements for each component in relationship Equation 1 above. There are currently no official estimates of consumption at the county level, so we will use the log of total income of individuals that reside in the county as the key left-hand side variable, which is independent of our right-hand and side estimates that we enter as log consumption. We believe that this is a reasonable proxy for consumption as about 90 percent of disposable income is spent on Personal Consumption Expenditures (PCE), but this does introduce elements in the accounting relationship that are missing on the right hand side (e.g., savings and taxes).

To account for these missing factors, we will focus on explaining the general relationship in terms of log income on the left-hand side and log consumption on the right hand side. We will examine this relationship in terms of levels, but also growth, where growth allows for the

Figure 1: Distribution of Net Exports both Unweighted and Weighted by Dollars Produced



inclusion of county-specific fixed effects. On the right hand side of the equation, we have estimates of final products sold to consumers in the county in each year, accounting for both total production in the county and net income. To demonstrate the importance of the geographic linkages alone, we use a single year's estimate of the credit card data (i.e., 2015) to estimate consumption flows across regions, assuming this flow relationship remains constant over time. We then investigate whether the inclusion of imports and exports helps improve the association between income (our proxy of consumption) and final products sold.

Specifically, we explore this analysis using the simple regression model of the following form:

$$\begin{aligned}\log(\text{Income}_{i,t}) &= \beta_1 \log(\text{Final Products Sold}_{i,t}) \\ &+ \beta_2(\log(\sum_{j\neq i} \text{Share Merchant Export}_{j\to i,T} \cdot \text{Final Products Sold}_{j,t}) \\ &+ \beta_3 \log(\sum_{j\neq i} \text{Share Merchant Export}_{i\to j,T} \cdot \text{Final Products Sold}_{i,t})) + \gamma_i + \varepsilon_{i,t} \qquad (3)\end{aligned}$$

where γ_i are county-specific fixed effects, which are included in some specifications. (Note that if fixed effects are included, then β_1 and β_3 cannot both be identified as they are both functions of Final Products Sold$_{i,t}$.) If imports and exports are unimportant then we should expect that final products sold in a county should have a close relationship to income in the county (our proxy to consumption), with little explanatory power from imports and exports. In addition, absent major changes in purchasing habits (e.g., constant budget shares) we should expect the coefficient on $\beta_1 + \beta_2 + \beta_3$ to be roughly 1, so that consumption changes roughly proportional to income.

Table 4 shows the relationship between income and the components of consumption in the county, including final consumption produced, consumption exported and consumption imported. The first two estimates exclude the county-specific fixed effects, so they are showing the basic cross-sectional relationship. Here we see that final goods produced in a county are closely related to final goods consumed, as expected, but this estimate essentially captures differences in the relative sizes of counties. The second column in the table includes the export

Table 4: Regression of Log(Income) on Receipts, Imports and Exports

	OLS	OLS	F.E.	F.E.
Log(Total Receipts)	0.981***	0.776***	0.323***	0.243***
	(0.00969)	(0.0371)	(0.0450)	(0.0288)
Log(Export)		-0.217***		
		(0.0280)		
Log(Import)		0.479***		0.838***
		(0.0451)		(0.275)
fixed effects	no	no	yes	yes
N	15312	15250	15312	15250
r2	0.977	0.985	0.794	0.809

Standard errors in parentheses

All estimates include state-level fixed effect clusters.

* p¡0.10, ** p¡0.05, *** p¡0.01

and import variables, which has the expected negative sign on exports and positive sign on imports, indicating that accounting for net-exports is important fo understanding consumption patterns in a cross section.

A more rigorous test of the importance of net-exports in consumption is shown in the panel estimates, that include for unique county-specific fixed effects and examine the relationship between growth in final products sold and income over the time period from 2012 to 2017. (Since we hold consumption flows fixed at 2015 levels we find exports and final goods produced are collinear, since exports are calculated by multiplying exports by final production.) If we do not account for imports in consumption we find an inelastic relationship between final products purchased and income, with an elasticity of just 0.32. However, we see that accounting for net imports into the region is an important component of the consumption measure and we are not able to reject a roughly unit-elastic relationship between income and consumption (0.88+0.24=1.12).

The relationship between these revenue shares and income provides an important validity check on the data constructed for the analysis.

7 Empirical Model of Consumption Flows

Recall that the main estimate considered here has the following function form:

$$\ln(S_{lj}) - \ln(S_{l0}) = t(l,j) + \beta_i \cdot x_j + \alpha \cdot p_j + Q_j + \xi_{l,j,n}$$

There are a number of details that must be addressed before bringing the model to the data: (1) defining the shares and the measurement of price; (2) defining the outside good; and (3) measurement error.

7.1 Defining shares and the measurement of price

The price variable used in the analysis is simply the geometric mean of the dollars spent per card swipe. However, the typical transaction could contain a variety of goods and services that are not comparable to NAICs in the same industry at another location. For instance, in one are of the country consumers may tend to buy 5 items per transaction (i.e. card swipe), while in another area it could be common to buy 4 . This would show up as a price difference acros areas. We address this issue by using an estimated average price based on spending coming across different areas. For instance, for the price paid for restaurants in Montgomery County, Maryland we will consider the geometric average price per dollar spent based on consumers in Montgomery County, but also the geometric average price paid by consumers coming from all other counties in the country. The averaging over individuals from different locations may help standardize the amount purchased per card transaction (e.g., wealthy individuals purchasing more items). Once the price is defined, we compute the estimated number of transactions as the total estimated revenue divided by the estimated average price.

Note that even for an accurately measured price variable, we anticipate the potential for endogeneity in this price estimate. We would expect that potential endogeneity would add a positive bias to the price coefficient, as areas with higher quality of consumption would likely be able to charge a higher price. For this reason, this method would likely undervalue higher quality areas. We will keep this in mind as we turn to the interpretation of the key estimates. In

addition, we plan to investigate alternative methods of construction, such as using regional price indexes constructed by the Bureau of Economic Analysis

7.2 Defining the outside good

To define the outside good, we first examine the distribution of the number of goods purchased per capita in each NAICs category across all counties in the United States. We then define the total potential market for a particular category to be determined by the 99th percentile of the units of consumption for a product. Those areas that consume more than the 99th percentile are potential outliers, so we remove those estimates from consideration.

Alternative approaches to define and measure the outside good will be employed to determine the robustness of our results to alternatiev empirical assumptions.

7.3 Measurement error

There are a number of steps taken to derive the share of consumption across areas. There may be measurement error for any particular area or industry potentially stemming from multiple causes: estimates of revenue flows, the measurement of price, or the estimates of receipts. This measurement error is a major empirical justification for focusing on multiple industries and an average component of quality, Q_j. If the factors that may create measurement error for one sector or location are uncorrelated with the measurement error for other sectors, aggregation will attenuate this error.

8 Empirical Results and Geography of Consumption

The above utility model is estimated with flexible third order polynomials of disutility from travel for every industry. Select coefficients from this model are reported in Table 5. In particular, the elasticity with respect to price across different industries. We also include an interaction between the home location of a consumer and industry, to capture the fact that consumption patterns may

be distinct for individuals purchasing in their home county, relative to consumers traveling to consume in that same county.

Table 5: Conditional Logit Regression on Across Area Consumption

	(1) meanutility	(2) meanutility	(3) meanutilityspend
lprice	-0.675***	-0.669***	0.327***
	(-27.33)	(-27.28)	(13.35)
ldistance	-1.274***		
	(-122.22)		
sim_race_index	-0.863***	-1.006***	-0.971***
	(-14.89)	(-18.83)	(-18.27)
sim_educ_index	-1.851***	-1.445***	-1.402***
	(-27.15)	(-21.09)	(-19.82)
Observations	25347658	25347658	25612841
Adjusted R^2	0.684	0.701	0.702

t statistics in parentheses
All estimates include county fixed effects.
* p¡0.10, ** p¡0.05, *** p¡0.01

9 Empirical Results

The quality measures Q_j are all produced at the county level. An MSA or non-MSA area of a state is often made up of multiple counties. To obtain a measure of average quality for an MSA we average over counties, we weight the quality of consumption in each county, Q_j, by the retail sales in that county, to measure the average quality per dollar of spending in the MSA. The average captures the average quality of spending in the MSA.

Table 6 shows the resulting rankings for the quality of consumption rankings. Several large cities are included among the top 25, including Los Angeles, Chicago, Seattle, and Dallas. The list is also full of potential vacation locations including Las Vegas, Miami, and Honolulu. The bottom rankings includes relatively rural cities such as Cumberland, Maryland and Weirton-Steubenville, West Virginia/Ohio. The highly ranked cities not just major tourist destinations. Salt Lake City and Tucson do well. It is also not the case that rural areas are always low ranked mechanically

Table 6: QOC Top 25 Areas

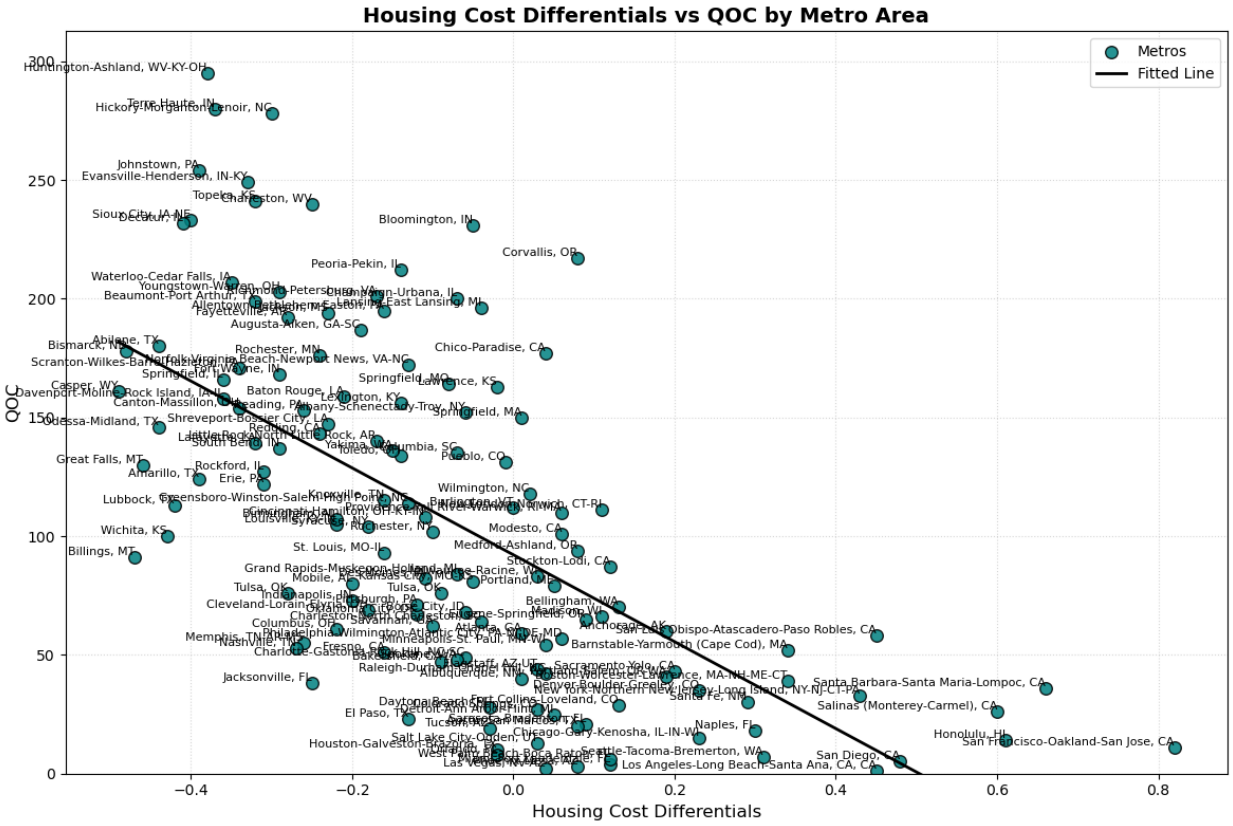
Area	Rank
Los Angeles-Long Beach-Santa Ana, CA	1
Las Vegas-Paradise, NV	2
Phoenix-Mesa-Scottsdale, AZ	3
Miami-Miami Beach-Kendall, FL	4
San Diego-Carlsbad-San Marcos, CA	5
West Palm Beach-Boca Raton-Boynton FL	6
Seattle-Bellevue-Everett, WA	7
Orlando, FL	8
Dallas-Plano-Irving, TX	9
Houston-Baytown-Sugar Land, TX	10
San Francisco-San Mateo-Redwood City, CA	11
Cape Coral-Fort Myers, FL	12
Salt Lake City, UT	13
Honolulu, HI	14
Chicago-Naperville-Joliet, IL	15
San Antonio, TX	16
Tampa-St. Petersburg-Clearwater, FL	17
Naples-Marco Island, FL	18
Tucson, AZ	19
Austin-Round Rock, TX	20
Sarasota-Bradenton-Venice, FL	21
Myrtle Beach-Conway-North Myrtle Beach,	22
El Paso, TX	23
HI Non-metro Area	24
Detroit-Livonia-Dearborn, MI	25

due to the low density of retail outlets. Rural Hawaii scores in the top 25. Non-MSA Vermont, California, and Colorado also score highly, above most urban areas.

Figure 2 reports the scatterplot of QOC and housing costs. There is a correlation between the two indicators. There are no high QOC cities with very low housing costs and vice versa. However, there is heterogeneity that shows the value of an independent measure. We find several low housing cost, high wage, and high QOC locations, including fast growing Texas cities like Dallas and Houston. Methods using residuals of wages and housing prices give them low rankings. They need a low amenity value to equalize locational value. Albouy (2019) finds very low amenity values for these cities. The high QOC ranking helps explain why they are growing.

We do a number of robustness checks on the rankings. First, we do a correlation with the Albouy (2019) rankings. The rankings are, if fact, distinct, as we should expect, but we find a

Figure 2: QOC and Housing Costs



strong positive correlation in the rankings using the Albouy adjusted ranking estimates, but we find a strong negative correlation using the unadjusted rankings.

10 Conclusion

We present a simple model of consumption preferences across locations, which we use to estimate "consumption amenities" of different locations across the U.S. We use novel data of credit card transactions to measure these amenities using discrete choice demand models, which show varied consumption patterns across industries. We show that the consumption amenities are highly correlated with housing prices and other measures of amenities. However, we also identify high quality of consumption areas that do poorly in the traditional methods. These locations are fast growing, suggesting they are attractive.

11 References

Auerbach, Alan, and Yuriy Gorodnichenko, (2012). "Measuring the Output Responses to Fiscal Policy," American Economic Journal: Economic Policy 4(2), May, 1-27.

Auerbach, Alan, Yuriy Gorodnichenko and Daniel Murphy. (2019) "Local Fiscal Multipliers and Fiscal Spillovers in the United States", Working Paper

Albouy, David, (2019), "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas." Working Paper

Batch, Andrea, Benjamin Bridgman, Abe Dunn, and Mahsa Gholizadeh, (2025) "Consumption zones", Journal of Economic Geography 25(2), pp. 191–213.

Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins. (2016) "A Dynamic Model of Demand for Houses and Neighborhoods." Econometrica. 84(3), pp. 893-942

Blanchard, Paul, Douglas Gollin, Martina Kirchberger, and Michael Peters (2025) "Human Mobility and Spatial Frictions in Three African Countries," mimeo, Tufts University.

Blomquist, Glenn C., Mark C. Berger, and John P Hoehn (1988) "New Estimates of Quality of Life in Urban Areas." American Economic Review, 78, pp.89-107.

Couture, Victor and Jessie Handbury (2020). Urban revival in America. Journal of Urban Economics 119: 103267.

Davis, Donald, Jonathan Dingel, Joan Monras, and Eduardo Morales (2019) "How Segregated is Urban Consumption?" Journal of Political Economy, 127(4), pp. 1684-1738.

Dolfen, Paul, Liran Einav, Peter J. Klenow, Benjamin Klopack, Jonathan D. Levin, Larry Levin, "Assessing the Gains from E-commerce." Working Paper.

Dunn, Abe, and Mahsa Gholizadeh (2025). "The Geography of Consumption and Local Economic Shocks: The Case of the Great Recession." American Economic Journal: Macroeconomics 17 (2): 206–44.

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

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. (2018), "Commuting, Migration, and Local Employment Elasticities," American Economic Review, 108(12): 3855–3890

Nakamura, Emi , Jón Steinsson (2014) "Fiscal Stimulus in a Monetary Union: Evidence from U.S. Regions" American Economic Review, 104(3), pp. 753-792.

Rappaport, Jordan (2008). "Consumption amenities and city population density" Regional Science and Urban Economics, 38 (6): 533–552.

Roback, Jennifer (1982) "Wages, Rents, and the Quality of Life." Journal of Political Economy, 90. pp. 1257-1278.

Rosen, Sherwin (1979) "Wages-based Indexes of Urban Quality of Life." in P. Mieszkowski and M. Straszheim, eds. Current Issues in Urban Economics, Baltimore: John Hopkins Univ. Press.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, (2019), "Diverging Trends in National and Local Concentration" Working Paper.

Syverson, Chad, (2019). "Macroeconomics and Market Power: Facts, Potential Explanations and Open Questions." Brookings Institution working paper.

12 Appendix

12.1 Steps for processing card transactions

The share of merchants transactions going to a particular location for a given industry is estimated based on a county-pair dummies along with distance across georaphies interacted with the associated NAICs industry, as well as other covariates. The county-pair dummy capture the typical transaction flows between two counties across all industries, with the distance interaction accounting for the feature of the data that individual are more likely to travel to consume for some industry categories relative to others. If an industry is not imputed based on this first imputation strategy, if, for example, there is no transaction between the counties in the data, then a second stage estimate is applied that includes county dummies along with various covariates, including distance terms interacted with the industry.

Before estimating our model of consumer demand, we first take steps to validate the observed consumption flows. In particular, the consumption flows across regions provide a unique connection between the production in regions with associated consumption, through a basic

accounting relationship that consumption is equal to production of final goods in an area minus net exports of consumption. While we do not have an independent estimate of consumption, we do have independent estimates of income that has been shown to be closely tied to consumption, as 90 percent of disposable income goes toward consumption. We show that the explanatory power of net exports is as large as the explanatory power of GDP in explaining changes in income over the 2012 to 2017 period, as would be suggested by the basic accounting relationship. These estimates show that the geographic connection between consumption and production is an essential component for understanding economic growth.

The FirstData spending flows across geographies are important more generally as they connect the production of final demand for many sectors to the consumption of final demand. This connection is important for a variety of purposes XXX. The location choice problem reveals many unique features about the geography of consumption in the United States. Unique insights ZZZ. In particular, we find that the geographic distance individuals are willing to travel and consume varies greatly across industries. Extremes ZZZ. The geographic areas in which different industries compete and provide services various greatly across industries. The per mile disutility for travel is highest for shorter distances, but lower for long distance, likely reflecting the lower costs as individuals change the mode of travel. Price is negatively correlated with consumption patterns, and we find it important to instrument for price to obtain reasonable elasticities.

Alternatively, the same model may be motivated through the trade-literature and XXX. demand model is also related to CES models of trade and demand across locations.

amenities across locations implementing this idea is that traditional data sources provide information on final sales to consumers (e.g., economic census) or information on incomes of consumers or consumption (e.g., CEX or XXX), but there is limited data connecting consumption locations to income. One area where this relationship is well studied given the widely studied claims data is the health care sector ZZZ.

(Maybe a useful side comment more for presentations: The case of Las Vegas may be the key exception to prove the rule, since it is a city marketed with the slogan: "what happens in Vegas stays in Vegas.")

For this reason, we focus on estimates where the demand for consumption is estimated jointly across multiple sectors. The basic idea is that factors that may create measurement error for one sector or location may be uncorrelated with the measurement error for other sectors.

The multiple industries are estimated jointly to capture a common quality component, Q_j, across industries. Let each NAICs code industry be denoted by n,then the combined estimation for these industries is denoted as:

$$\ln(S_{l,n,j}) - \ln(S_{l,n,0}) = t_n(l,j) + \beta_n \cdot x_{j,n} + \alpha_n \cdot p_{j,n} + Q_j + \xi_{l,j,n}$$

There is also a theoretical justification for estimating the consumption jointly, as there are likely agglomorations in consumption, so we might expect there to be a common factor of consumption that we could estimate more precisely by combining industries. One natural source of agglomeration economies from consumption is that once a consumer pays the fixed cost to travel and consume in one location for one industry, the marginal transportation costs for additional consumption in the same area are likely minimal. The empirical justification is that there may be measurement errors in any one sector, potentially stemming from multiple causes: estimates of revenue flows, the measurement of price, or the estimates of receipts. However, the factors that may create measuement error for one sector or location may be uncorrelated with the measurement error for other sectors. We focus on an aggregate quality of consumption, Q_j.

Issues arise related to measuring the price, defining the outside good and measurement error problems that are unique to the data applied here. We take steps to overcome each of these issues. However, even with these potential limitations, we believe the main contribution of the paper is to present a new look at the geography of consumption, which has not previously been examined, which we hope opens up new avenues of future research and a broader understanding of the linkages among markets through consumption.