NBER WORKING PAPER SERIES

THE GEOGRAPHY OF CONSUMPTION

Sumit Agarwal
J. Bradford Jensen
Ferdinando Monte

Working Paper 23616
http://www.nber.org/papers/w23616

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2017

We would like to thank Steve Redding, Andres Rodriguez-Clare, Esteban Rossi-Hansberg, and Tony Yezer for helpful comments and discussions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w23616.ack

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

The Geography of Consumption
Sumit Agarwal, J. Bradford Jensen, and Ferdinando Monte
NBER Working Paper No. 23616
July 2017
JEL No. F1,F14,L8,R1,R2

ABSTRACT

We use detailed information from U.S. consumers' credit card purchases to provide the first large-scale description of the geography of consumption. We find that consumers' mobility is quite limited and document significant heterogeneity in the importance of gravity across sectors. We develop a simple model of consumer behavior, emphasizing the role of the durability/storability of products, to organize the main stylized facts. Heterogeneity in the storability of products across sectors generates a positive correlation between the strength of gravity and the frequency of transactions at the sector level; this correlation is a clear feature of the data. Using daily rain precipitation from thousands of weather stations in U.S., we show that shocks to travel costs change the spatial distribution of expenditure, and they do so differentially across sectors: hence, the level and heterogeneity of travel costs shape the level and elasticity of any merchant's demand. This evidence suggests that incorporating the demand-side is essential to analyzing the distributional consequences of local and aggregate shocks across regions. These results also suggest the demand-side is critical to understanding the location of firms and employment in the large and understudied service sector.

Sumit Agarwal
McDonough School of Business
Room 427
Georgetown University
3700 O Street NW
Washington, DC 20057
ushakri@yahoo.com

J. Bradford Jensen
McDonough School of Business
Georgetown University
Washington, DC 20057
and Peterson Institute for International Economics
and also NBER
jbj24@georgetown.edu

Ferdinando Monte
Georgetown University
McDonough School of Business
37th & O Streets, NW
Washington, DC 20057
ferdinando.monte@georgetown.edu

1 Introduction

The equilibrium production and location decisions of firms depends, among other things, on the characteristics of local demand and, in particular, on the willingness of consumers to travel to buy goods and services. However, lack of direct evidence on consumers' mobility has limited our ability to characterize local consumption markets. In this paper, we focus on the consumers' demand side and provide the first large scale description of the geographical dimension of different consumption markets using more than 1.7 million individual American consumers' credit card transactions.

We organize our analysis around the impact of spatial frictions in the form of distance, since final consumption of goods and services typically requires a physical travel of one of the two sides towards the other. The data reveals that consumers typically visit just a few of the many available locations for purchases of goods and services. Moreover, as broadly documented for merchandise trade both at international and intra-national level, gravity is a first-order feature of the data: total expenditure of one place's residents on another place's firms declines with distance. To understand the sources of this decline, we decompose total expenditure in contributions coming from 1) average value of expenditure per account, and 2) total number of accounts; the average value of expenditure per account can be further decomposed into 3) the average number of transactions per account and 4) the average value of a transaction. We analyze the spatial decay of each component separately.

We find large heterogeneity across industries in the overall impact of distance and in the importance of extensive margins (that is, in how much of the decline is accounted for by fewer people spending, or by people spending less frequently). We also find that the differences in gravity across industries are correlated with the frequency of transactions. We relate this frequency to the durability/storability of the final item demanded, which we interpret as a new determinant of gravity. While economists have long studied how exogenous characteristics of an industry (input-intensity or the weight-to-value ratio) have influenced trade patterns, most research has implicitly assumed a binary representation of durability/storability: goods are durable/storable and services are not, and thus, goods are tradable and services are not. By examining expenditure patterns at a small geographical scale, we are able to uncover a more nuanced picture related to the durability/storability of the good.

We set up a highly stylized model of consumers' cost minimization to rationalize these findings. In a given sector, varieties can be purchased at home or outside at different prices. Consumers would prefer to travel if they could buy varieties at a cheaper price. Travel is expensive, which creates an incentive to minimize trips; storage is also expensive, however, and this induces a trade-off between number of trips and purchase size. Consumers in our model choose which varieties to purchase at home vs. outside, how many trips to take and the "batch size" of each purchase. This simple model is able to organize our stylized facts. Gravity emerges since, as travel costs increase, those consumers who travel outside will buy fewer varieties (intensive margin), and some of those will stop traveling altogether (extensive margin). Moreover, when a sector has higher storage costs (a simple way of modeling less durability), consumers respond by increasing the average frequency of trips, and by traveling less outside: this response reproduces the cross-sectoral correlation between stronger gravity and higher trips' frequency.

While these cross-sectional facts are interesting and suggestive of an active role of consumers' travel cost in shaping local demand, unobserved heterogeneity in sector characteristics might be a contributing factor in determining the observed strength of gravity. To provide more direct evidence on the role of demand, we use exogenous variation in consumers' travel costs induced by rainfall. We use daily data on rain precipitation from thousands of weather stations to examine the response of expenditure to higher travel costs. We find that rain reduces expenditure of consumers living in a particular place both at home and in outside locations. Importantly, we also find that the spatial distribution of demand changes in a reasonable way: rain induces consumers most sensitive to travel costs not to go out and spend, so that only those less sensitive to travel costs are observed purchasing. This makes gravity flatter. From the perspective of a merchant located in a particular place, the level and heterogeneity in travel costs will have an impact on the level and elasticity of demand.

Understanding the sources of gravity has important implications for defining the spatial dimension of a market and for the response of local outcomes to local exogenous shocks. In particular, since gravity varies across sectors, our results inform how demand considerations shape the spatial distribution of production in the large (and understudied) service sector. Final consumption accounted for around 70% of GDP in 2015 in the United States; the service industries involved in its delivery, from apparel stores, to restaurants and personal services providers, accounted for around 70% of employment, and more than 80% of total value added. Understanding the spatial determinants of consumption is therefore essential for a wide spectrum of issues: from the degree of spatial competition between firms to the determination of local and aggregate productivity and factors' income, from the consequences of local labor demand shocks, local taxes and regulation, to the impact of investment in transportation infrastructure and of other "place-based policies". Further, our results provide important information for the study of the liberalization of international trade and investment in services, since the extent to which foreign direct investment flows to a country (and to which part of the country) depends crucially on how local the market for a particular service is.

While spatial analyses of the manufacturing sector abound, they are not particularly informative regarding the final consumer behavior. The practical importance of direct sales from manufacturers to consumers is still somewhat limited: for example, e-commerce sales (which include sales from a company's website, but also indirect sales from other distributors) account for only 6.4% of total retail sales in 2014, and only 0.9% in 2000, closer to our sample period, 2003 (Hortaçsu and Syverson, 2015). Intranational surveys on goods' flows typically record firm-to-firm transactions. The limited literature on spatial competition in consumption markets mostly focuses on specific industries. For the restaurant industry, Couture (2016) evaluates the extent to which consumers gain from increased density and finds that larger variety (rather than lower travel times) are the main driver of consumers' valuation of higher density; Davis, Dingel, Monras and Morales (2016) discuss the relative impact of spatial vs. social friction in restaurant consumption. In the food distribution sector, Handbury, Rahkovsky and Schnell (2016) study the role of spatial access to healthy food supply in explaining differences in the quality of food intake across income groups, and argue that observed differences in access are most plausibly the result of optimal supply responses to differences in demand. Other industries that have been studied include gasoline (Houde, 2012) and movie theaters (Davis 2006). Chandra, Head and Tappata (2014) find that consumers' travel across the U.S.–Canada border responds to real exchange rate movements and distance to the border; they propose and estimate a simple model to rationalize those findings. We differ in that we emphasize travel and storage costs, which allows us to study cross-sectoral differences in the strength of gravity and the relation with the frequency of transactions; we also have the ability to explore a richer geography and study expenditure levels, rather than simple trip counts. Studies focusing on consumption across cities, with less focus on consumers' mobility, include Glaeser, Kolko and Saiz (2000) (who explore the increasing importance of cities as consumption centers) and Schiff (2015), who finds that larger and denser cities offer more restaurant varieties. Overall, we contribute to this literature by providing results which are comparable across industries, extend the set of industries for which we can assess consumers' mobility, and exploit cross-industry variation to argue that storability of a good provides a new determinant of gravity. We do so by building on and extending the literature on spatial frictions (Anderson 1979, Anderson and Van Wincoop 2003, Eaton and Kortum 2003, Hummels and Klenow 2005, Hillberry and Hummels 2008).

The retail sector is also subject of a growing literature. Jarmin, Klimek and Miranda (2005) and Hortaçsu and Syverson (2015) present some overall trends in the industry. The important role of the retail sector in price determination is emphasized by Nakamura (2008), who shows that a majority of retail-store price variation is attributable to retail chain–level shocks, while only a minor fraction to manufacturers or wholesalers–specific shocks; this emphasizes the importance of understanding pricing (and hence demand) conditions at the retail level. Bernard, Jensen, Redding and Schott (2010) examine characteristics of wholesalers and retailers involved in international transactions, finding that they are significantly smaller compared to their "producer and consumer" counterparts.[1]

Our work is also relevant for the growing literature on e-commerce and on-line transactions: firms operating in this way are competing with brick-and-mortar stores precisely taking into account consumers' travel costs. Aspects of on-line vs. off-line retail are analyzed in Ellison and Ellison (2009), who study the importance of taxes in determining sales of on-line versus traditional retailers: among other things, they find that geography still matters (consumers prefer to buy from home state or neighboring retailers after accounting for other factors), albeit the effect of proximity via shipping times is small. The importance of distance and the persistence of a home bias is also found in on-line auctions (Hortaçsu, Martínez-Jerez and Douglas 2009). Einav et al. (2017) use credit card transactions to quantify the gains from e-commerce.

We proceed in the paper by presenting in Section 2 a simple model of shopping behavior; this model will introduce concepts and terms which will organize our main cross-sector stylized facts in Section 3. In Section 4 we turn to a within-sector analysis and show how the spatial distribution of expenditure responds to rainfall shocks. Sector 5 concludes.

[1] To our knowledge, the first formulation of a gravity law in the retail sector dates back to Reilly (1931): "Two cities attract retail trade from any intermediate city or town in the vicinity of the breaking point, approximately in direct proportion to the populations of the two cities and in inverse proportion to the square of the distance from these two cities to the intermediate town."

2 A Simple Model of Shopping

We start by developing a highly stylized model of consumer behavior to help guide our empirical analysis and introduce some terminology. The model will only study the cross-sectional implications of cost-minimizing consumers. We combine ideas present in Dornbush, Fisher and Samuelson (1977) and Oi (1992).

2.1 Setup

Let us focus our attention on a world with two locations, h and s (home and an additional outside sale location). These locations are identical: we eliminate the role of productivity differences and market size to focus solely on the consequences of travel costs.[2] There is one sector, with a unit continuum of varieties, $j \in [0, 1]$. A mass 1 of consumers, indexed by ω, wants to buy a fixed quantity $\bar{q}$ of each variety. Consumers differ in the travel costs faced to buy varieties in the outside location. In particular, a consumer ω faces a cost of a trip outside $t_s(\omega)$; all consumers face the same positive cost of a trip home, $t_h \leq \min_\omega t_s(\omega)$. It will turn out below that a useful summary of consumers' travel cost is $t(\omega) \equiv t_s(\omega)^{1/2} - t_h^{1/2}$; for now, we introduce the notation $T(.)$ for the CDF of this cost difference across consumers, and note that it has positive support.[3] We assume that the travel cost to buy a measure $\bar{j}$ of varieties in a location n is given by $\bar{j} \times t_n(\omega)$: this a natural assumption, as it says that for a travel to a given location, it takes more time to buy several varieties. Consumers face a storage cost g for each of the varieties. The trade-off between storage and travel costs shapes consumers' decisions of where to buy each variety, how frequently, and how much to purchase per trip.

Each variety j is potentially sold in two places, and so it has two shop prices, $p_h(j)$ and $p_s(j)$. We order the varieties j so that the difference

$$d(j) \equiv p_h(j) - p_s(j) \tag{1}$$

is increasing in j: when j is higher, residents in h would find it increasingly convenient to buy a single variety in location s. We assume $d(0) < 0$ and $d(1) > 0$ and define

$$j_z : d(j_z) = 0 \tag{2}$$

as the variety whose price is the same in both locations. Our initial assumption of identical locations implies that $j_z = 1/2$, and $-d(j - x) = d(j + x)$ for $x \in [0, 1/2]$ (that is, the branch of $d(j)$ to the right

[2] In our empirical analysis we will always include residence and shopping location fixed effects, typical in the trade literature. Our assumption of identical locations maps consistently into this empirical strategy.

[3] Most of the results below can be generalized to an environment where consumers are described by a continuous bivariate distribution over costs of travel at home and outside, $\{t_h(\omega), t_s(\omega)\}$ with $t_h(\omega) \leq t_s(\omega)$. This generalization is possible exactly because for most results consumers will only act upon $t_s(\omega)^{1/2} - t_h(\omega)^{1/2}$.

of $j = 1/2$ is the reflection of the branch to the left along the j-axis). Moreover, we denote with

$$C_n(j', j'') \equiv \int_{j'}^{j''} \bar{q} p_n(j)\, dj \tag{3}$$

the cost of buying a range of varieties from j' to j'' in location $n \in \{h, s\}$. We now turn to an exploration of the optimal behavior of a cost-minimizing consumer. In the following subsection, we will then explore the cross-sectional implications of such behavior.

2.2 Consumer behavior

A consumer ω wants to minimize the cost of purchasing $\bar{q}$ units of each variety, by choosing the location(s) of where to buy, the range of varieties bought from each location, and the size of the purchase for each variety (i.e., how frequently to go to the store).

We can immediately exclude that consumers buy all goods in location s since $t_h < t_s(\omega)$ and $d(j) < 0$ for j low enough: there are always some varieties that are cheaper at home, and home is cheaper to reach. Hence, the consumer is choosing whether to buy only in h, or in both h and s, and which goods to buy where, in the second case. Since consumers are always buying something at home, the range of goods bought in s will take the form $[\bar{j}, 1]$ (i.e., they will buy in s some of the cheapest varieties). Consumers choose a "batch size" z_n for every trip, so that $\bar{q}/z_n$ is the number of trips to location n's shops, and $z_n/2$ is the average inventory of each variety bought from location n (we will ignore integer constraints). Consumers pay a holding cost g per unit of average inventory of each variety, so that holding a higher average stock of any given variety is more expensive.

Given these considerations, the consumer problem can be formally expressed as:

$$C(\omega) = \min_{\substack{I\in\{0,1\},\bar{j}\in[0,1],\\ z_h^{(I)}\in[0,\bar{q}],z_s^{(I)}\in[0,\bar{q}]}} (1-I) \times C^{(I=0)}(\omega) + I \times C^{(I=1)}(\omega) \tag{4}$$

where

$$C^{(I=0)}(\omega) \equiv C_h(0,1) + 1 \times \left[t_h \frac{\bar{q}}{z_h^{(I=0)}} + g \frac{z_h^{(I=0)}}{2} \right] \tag{5}$$

$$\begin{aligned} C^{(I=1)}(\omega) \equiv\; & C_h(0,\bar{j}) + \left[t_h \frac{\bar{q}}{z_h^{(I=1)}} + g \frac{z_h^{(I=1)}}{2} \right] \times \bar{j} + \\ & + C_s(\bar{j},1) + \left[t_s(\omega) \frac{\bar{q}}{z_s^{(I=1)}} + g \frac{z_s^{(I=1)}}{2} \right] \times (1-\bar{j}) \end{aligned} \tag{6}$$

Equation (4) says that the consumer will choose the minimum cost of procuring the desired bundle of varieties by choosing whether to buy only at home ($I = 0$) or at home and travel outside ($I = 1$); when $I = 1$ we will say that "the consumer travels," for brevity. Equation (5) gives the cost of buying the whole range of varieties from home, accounting for travel costs to buy a measure 1 of varieties in batches of size $z_h^{(I=0)}$: this will require $\bar{q}/z_h^{(I=0)}$ trips and imply storage costs of $g z_h^{(I=0)}/2$ per variety. Equation

(6) gives the cost of buying a range of varieties $[0,\bar{j}]$ at home, and a range $[\bar{j},1]$ of varieties outside: the travel and storage costs of batches $z_h^{(I=1)}$ and $z_s^{(I=1)}$ per variety are now multiplied by the correspondent measure of varieties bought in each location.

Suppose first the consumer chooses to buy everything at home. Then, $\bar{j}=1$, and

$$z_h^{(I=0)} = \left(\frac{2\bar{q}}{g}t_h\right)^{1/2} \tag{7}$$

$$f_h^{(I=0)} = \left(\frac{g\bar{q}}{2}\frac{1}{t_h}\right)^{1/2} \tag{8}$$

are the batch size and the frequency of purchase. A consumer ω will buy larger batches to compensate for the extra cost of storage with savings from less frequent trips. In optimum, when travel costs are higher, the optimal batch is larger and the number of trips is lower. The total cost of purchasing all the units from home is then

$$C^{(I=0)}(\omega) = C_h(0,1) + \gamma\,(g\bar{q}t_h)^{1/2} \tag{9}$$

with $\gamma \equiv 2^{-1/2}+2^{1/2}$ a constant.

Now, suppose instead the consumer chooses to buy in both locations. The optimal batch sizes have an analogous functional form, but we now have two optimal batches, one for home and one for outside purchases:

$$z_n^{(I=1)} = \left(\frac{2\bar{q}}{g}t_n(\omega)\right)^{1/2},\ n\in\{h,s\} \tag{10}$$

Since the travel cost to s is higher, consumers will buy larger batches per trip outside, and make fewer trips with respect to home purchases: the average number of trips if a consumer travels is

$$f^{(I=1)} = \frac{1}{2}\left(\frac{g\bar{q}}{2}\frac{1}{t_h}\right)^{1/2} + \frac{1}{2}\left(\frac{g\bar{q}}{2}\frac{1}{t_s(\omega)}\right)^{1/2} \tag{11}$$

which is lower than (8). Which varieties would this consumer buy from home? Interestingly, a cost-minimizing consumer would buy some of the overpriced varieties from home. The first order condition with respect to the optimal variety to buy requires

$$p_h(j^*) - p_s(j^*) \equiv d(j^*) = \frac{g}{2}\left(z_s^{(I=1)} - z_h^{(I=1)}\right) + \left(t_s(\omega)\frac{\bar{q}}{z_s^{(I=1)}} - t_h\frac{\bar{q}}{z_h^{(I=1)}}\right) \tag{12}$$

This condition is very intuitive: it says that savings on the marginal variety bought outside must exactly compensate the extra storage and travel costs. If travel costs at home and outside were the same, batch sizes would also coincide and the consumer would buy each variety in the cheapest location. If the right-hand side is positive (as it turns out to be since $t_s(\omega) > t_h$), then the consumer is willing to buy some of the (least) overpriced varieties at home: this makes him spend a little more on some varieties, but he does not need to pay the higher travel costs and carry the much larger inventory necessary for them if they were bought outside.

Substituting the optimal batch size in (12), we obtain:

$$j^*(\omega) \quad : \quad \min\left\{d(j^*) = \left(\frac{\bar{q}g}{2}\right)^{1/2} t(\omega), 1\right\} \tag{13}$$

$$t(\omega) \quad \equiv \quad t_s(\omega)^{1/2} - t_h^{1/2} \tag{14}$$

Since $t(\omega) > 0$, $d(j^*) > 0$ and hence there is a unique optimal variety $j^* > 1/2$ for any given (differences in) travel costs, $t(\omega)$: even with identical locations, consumers will buy more varieties from home than from the outside location. When the difference $t(\omega)$ is above a threshold, the consumer finds himself choosing $j^*(\omega) = 1$ anyway; we call this threshold level of the difference in travel costs $t_{\max} \equiv d(1)(\bar{q}g/2)^{-1/2}$. The left panel of Figure 1 shows the optimal j^* as a function of t.

The minimum cost to buy varieties in both places is then

$$C^{(I=1)}(\omega) = C_h(0, j^*(\omega)) + C_s(j^*(\omega), 1) + \gamma(g\bar{q})^{1/2}\left[t_h^{1/2} j^*(\omega) + t_s(\omega)^{1/2}(1 - j^*(\omega))\right]$$

Consumer ω will choose to purchase in both locations if $C^{(I=1)}(\omega) \leq C^{(I=0)}(\omega)$, that is, if

$$\gamma(g\bar{q})^{1/2} t(\omega) \leq \frac{C_h(j^*(\omega), 1) - C_s(j^*(\omega), 1)}{1 - j^*(\omega)} \equiv \frac{\bar{q}\int_{j^*(\omega)}^{1} d(j)\, dj}{1 - j^*(\omega)} \tag{15}$$

which is again a very intuitive expression. The left-hand side is the net increase in travel and storage costs per variety due to the trip outside; the right-hand side represents the average saving per variety purchased outside. A consumer will travel if the savings generated by the trip outweigh those increases in costs. Note that what matters for the consumer choice of traveling is the summary difference $t(\omega)$, i.e., both (13) and (15) only depend on transportation costs through (14). In Appendix A we prove the following:

Proposition 1 *If $d(j)$ is strictly concave for $j \geq 1/2$, there exist a unique level of differences in travel costs t_c such that a consumer ω travels if and only if $t(\omega) \leq t_c$.*

If the only source of heterogeneity in ω is heterogeneity in t, we can write eq. (13) and (15) simply as a function of t. Figure 1 shows these two determinants of consumer choices as a function of t. The left panel shows the cutoff variety j^* as a function of travel costs t, as determined by eq. (13): if a consumer was constrained to travel, the marginal variety bought at home would increase with t (up to the point where a corner solution emerges). The right panel shows average savings and average costs per variety bought outside as a function of an individual consumer's travel cost; when average savings are above average travel and storage costs (green area) the consumer will choose to travel. This choice will be made only by consumers with a low enough travel cost (i.e., only for t low enough).

cutoff variety $j^*(t)$ cutoff travel cost t_c

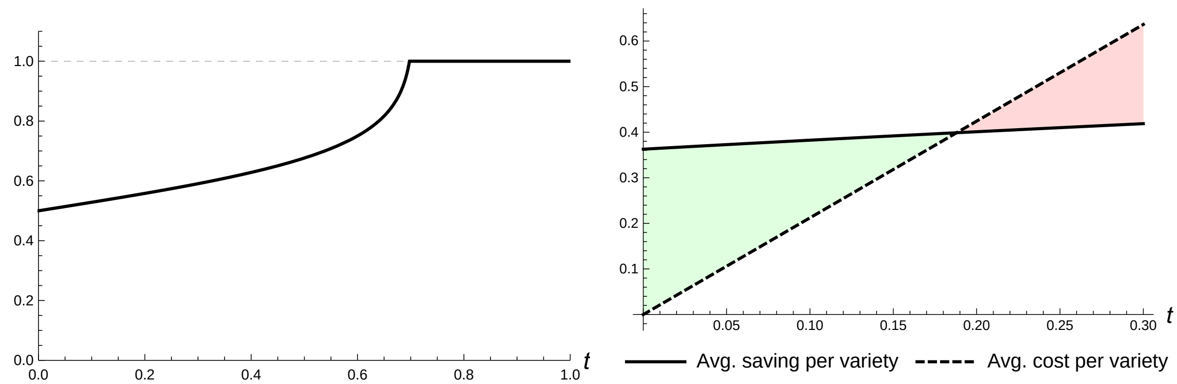

Figure 1: Consumers' choices

2.3 Cross-sectional implications

Having characterized the behavior of a single consumer, we now study the implications in a population of consumers heterogeneous in t, with CDF $T(\cdot)$. From Proposition 1 it follows immediately that the fraction of consumers who choose to travel is $T(t_c)$; hence, the total expenditure of consumers is

$$X = (1 - T(t_c)) C_h(0, 1) + \int_0^{t_c} C_h(0, j^*(t)) + C_s(j^*(t), 1) \, dT(t) \tag{16}$$

and the expenditure on the outside location is

$$X_{hs} = \int_0^{t_c} C_s(j^*(t), 1) \, dT(t) \tag{17}$$

We now want to understand the consequences of an increase in distance between h and s. In this stylized model of consumer behavior, this is equivalent to comparing an economy in which the distribution of travel costs is $T(t)$ to one in which the new CDF $T'(t)$ is shifted down and to the right. A simple way to achieve this shift is to assume that the travel costs of all consumers increase from t to κt, with $\kappa > 1$: this shift implies that $T'(t) = T(t/\kappa)$.[4]

We start by noting that absent general equilibrium effects, the analysis in the section above also describes the reaction of a consumer to increases in travel costs. If the travel cost of a consumer increases by κ, he chooses to buy fewer varieties outside (since $j^*(\kappa t) \geq j^*(t)$), so that the sectoral expenditure per consumer on varieties in s falls with higher distance, and X_{hs} declines. We will refer to this intensive margin as the "expenditure margin" below. In addition, for consumers such that $t < t_c < \kappa t$, the expenditure on outside goods drops to zero altogether: this reduction in the expenditure X_{hs} via a reduction in consumers traveling will be referred to as the "account margin" below. Overall, both forces push towards a reduction in expenditure on the outside location and an increase in the expenditure on the home location, i.e. expenditure patterns display gravity.

[4]Formally, T first-order stochastically dominate T', since $T(t) \leq T'(t)$ $\forall t$.

2.4 Cross-sector implications

We now compare the behavior of consumers as sector characteristics vary. This comparison is useful as we will find large heterogeneity in the strength of gravity across sectors, and a strong relation between gravity and the frequency of transactions.

Our stylized model naturally suggests differences in the costs of storage g as a possible source of heterogeneity across sectors. We think of this variable as a portmanteau that captures differences in shelf life, physical size, durability and storability across broad categories of goods. We start by establishing the following:

Proposition 2 *If the storage costs g increase, the travel cutoff t_c decreases, i.e., $dt_c/dg < 0$.*

which again we prove in Appendix A. An increase in storage costs requires that every consumer's savings on the first variety bought outside be higher (eq. (13)): varieties whose savings are too limited are now bought at home, so that every consumer buying in two locations will buy more varieties at home and fewer outside (the total expenditure X_{hs} decreases at the "expenditure" margin). Consumers that were already buying almost everything at home stop traveling altogether, which implies that the new indifferent consumer must be now one with lower travel costs: hence, t_c falls, and X_{hs} falls at the extensive "account" margin as well. Overall, sectors with high storage costs display stronger decline of expenditure as we move from home to the outside location, i.e. stronger gravity.

Another consequence of high storage costs is that the average frequency of trips increases. Define the population average frequency of trips as

$$\bar{f} = (1 - T(t_c))\, \bar{f}^{(I=0)} + T(t_c)\, \bar{f}^{(I=1)}$$

where $\bar{f}^{(I)}$ are the average frequency of trips in the populations of consumers whom choose to travel and not to travel. These average frequencies can be computed with appropriate integration using (8) and (11). In Appendix A, we prove the following:

Proposition 3 *If storage costs increase, the population average frequency of trips increases.*

As storage costs increase, individual agents optimally choose to reduce the batch size and increase the frequency of trips (immediate from (8) and (11)); hence both $\bar{f}^{(I=0)}$ and $\bar{f}^{(I=1)}$ grow. Moreover, since $T(t_c)$ falls, more weight is put on those who have a higher frequency of trips (those whom do not travel). Hence, the average frequency of trips in the population grows. The last two Propositions imply that, comparing different sectors, goods with high storage costs display stronger gravity and at the same time higher frequency of purchase.

Armed with these intuitions, we now start our exploration of consumers' geographic mobility.

3 The Geography of Consumption

3.1 Data description

We use a large proprietary dataset containing credit card transactions from a major financial institution. These transactions occurred roughly between February and October 2003. A transaction record contains, among other things, an exact date, an account ID, the amount spent, a Merchant Category Code (we will refer to it as a "sector") and information (to be processed) on the location of the merchant. In addition to all distinct transactions, we have information on the account itself; importantly, we know the ZIP code associated to it. The same data has been used in Agarwal, Marwell, and McGranahan (forthcoming). After cleaning procedures, we have 1,751,067 transactions for 71,927 accounts (see Appendix B for a complete description of the data cleaning and processing). The average transaction is 70 dollars, and total purchases amount to around $122 million. Table 1 reports a breakdown by 24 broad categories. The largest categories in terms of observations are Gasoline Services, Food Stores, Miscellaneous Retail, and Eating and Drinking Places. Table C.1 in the Appendix (page 36) shows summary statistics by State of purchase. The largest number of transactions are reported in New York, California, and Massachusetts.

In the remainder of this section we establish a number of stylized facts on the local nature of consumption markets. This analysis will be inherently cross-sectional in nature. In the following section we will study the consequences of a cost shock to show that the spatial distribution of expenditure reacts to consumers' travel costs for given location of consumers and merchants.

3.2 Consumers visit few locations

We start our exploration by considering how far consumers travel across locations for purchases. A "location" in the data is identified at the level of Census incorporated place or county subdivision. The data identifies expenditure flows between 17,572 unique locations (11,474 unique residence and 14,997 unique sale locations). The data records transactions among 238,269 unique pairs, 0.077% of all possible pairs. There are 7.3 transactions per pair, and the median pair has 1 transaction. Naturally, it is unrealistic to expect consumers to travel very long distances: for example, 2,588,158 potential location pairs (about 0.83% of all pairs) in our data have distance below 120 km.[5] Among those pairs, expenditure flows are recorded between only 122,130 pairs, or 4.7% of the possible pairs, with 14.3 transactions per pair on average, and 2 transactions for the median pair. Overall, the matrix of residence-sales location purchases is sparse.

Table 2 digs deeper into this limited consumer mobility. Its first row shows that consumers in the median residence visit only 6 distinct sales locations overall during the sample period (10.6 sales locations on average). One might think that this low number is simply a consequence of absence of close-by options. This is not the case. The second row in the Table shows that consumers living in the median residence

[5]Distance is always computed between the centroids of two locations using the Haversine formula. When looking at the impact of distance on flows below, we will also restrict our attention to transactions with distance up to 120km. Monte, Redding and Rossi-Hansberg (2016) find this threshold to be one where gravity in home-to-work commuting flows has a structural break, and seems to be a natural cutoff to focus on.

Table 1: **Summary of transaction amounts (in USD), by sector**

Broad Category	Median	Mean	St. Dev.	Sum	N
Agricultural Services	83	136	212	1,307,704	9,616
Amusement, Rec. Serv.	45	89	169	1,772,753	19,912
Apparel	49	75	114	6,117,925	81,846
Auto Repair/Service/Parking	41	151	325	3,465,225	23,009
Auto and Truck Sales/Service/Parts	66	198	423	6,625,779	33,486
Building Mat./Hardware/Garden Supp.	42	101	258	9,662,827	95,641
Communications	53	91	122	559,309	6,115
Durable Goods	68	209	520	839,482	4,026
Eating and Drinking Places	26	39	73	8,785,594	228,006
Educational Services	92	298	625	2,296,223	7,711
Food Stores	30	46	59	12,126,085	266,030
Furniture, Home Furnishings, Equip.	60	194	430	10,865,485	55,999
Gasoline Services	19	22	31	6,938,938	312,873
General Merchandise Stores	43	67	122	13,971,247	207,985
Health Services	71	164	375	4,490,768	27,403
Hospitality	96	170	308	6,436,342	37,971
Misc. Retail	32	65	182	16,113,793	248,288
Misc. Services	95	316	702	1,873,105	5,927
Motion Pictures	14	19	44	273,520	14,080
NonDurable Goods	38	78	175	640,316	8,249
Other Vehicles Sales/Service/Parts	76	259	746	1,366,524	5,280
Personal Services	37	74	212	2,417,176	32,575
Transportation Services	38	117	384	1,128,331	9,621
Vehicle Rental	125	189	230	1,778,716	9,418
Total	30	70	195	121,853,167	1,751,067

also have 186 sales locations within 120 km; the third row shows that overall the fraction of available locations where purchases actually occur is very small.[6]

What accounts for this low mobility? Overall, the number of visited locations grows at about half the pace of the available locations. In column 1 of Table 3 we regress the log of number of sales locations visited on the log number of sales locations available and find an elasticity of 0.54 (hence, well below 1). Distance on the other hand has a stronger role: controlling for the number of available sales locations, a 1% increase in average distance to those locations is accompanied by a 2.5% decrease in the number of locations visited (column 3). We now turn to a deeper exploration of distance.

[6] In Appendix C.2, page 35, we further show that this low mobility is not driven by accounts with low credit card usage overall.

Table 2: **Summary statistics across residence locations**

variable	min	p10	p25	p50	p75	p90	max	mean	N
Sales locations visited	1	1	3	6	13	26	447	10.92	11,179
Sales locations within 120km	2	61	109	186	327	636	1,116	265.73	11,179
Share available loc. visited	0	0.01	0.02	0.04	0.07	0.13	0.68	0.06	11,179

Table 3: Locations available and locations visited

Dependent variable:	Log of number of sales locations visited		
	(1)	(2)	(3)
Sales locations within 120km, log	0.539***		0.561***
	(0.010)		(0.009)
Average distance to sales locations within 120km, log		-2.361***	-2.594***
		(0.074)	(0.064)
Constant	-0.957***	12.011***	10.097***
	(0.051)	(0.319)	(0.276)
R^2	0.22	0.08	0.32
N	11,179	11,179	11,179

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

3.3 The distance traveled varies by sector

This first snapshot paints a picture where consumers have many options but choose to shop only in a limited number of locations; furthermore, a strong role is played by distance. A natural question to ask then is: how much do people travel for their purchases?

The median transaction in the data occurs at about 9 km from home. There is large dispersion around this typical value: the first 25% of transactions occur within the same place, while the third quartile is around 30 km. A long right tail of high distances is likely due to account holders traveling outside town for work or tourism.[7] While these and other details are relegated to the Appendix, we show in Figure 2 select percentiles of the distances at which transactions occur, by sector.[8] The heterogeneity in distance traveled is very significant: moving from a sector at the 10th percentile to a sector at the 90th, the median distance traveled goes up by a factor of around 7. The patterns make sense overall: the median transaction occurs at 4 km for staple items like Food Stores, and around 12 km for Eating and Drinking Places; it is, however, above 20 km for Durable Goods and 33 km for Amusement and Recreational Services, which

[7] Online transactions have been eliminated as much as possible. See the Data Processing section in the Appendix for more details.

[8] Tables C.3 and C.4, in pages 38 and 39 respectively, show percentiles in the distribution of transaction distances by sector in the raw data and weighted by value of the transaction, respectively. The typical dollar is spent further than where the typical transaction occurs, as reflected in right-ward shifts in the value-weighted distributions.

are likely purchased less frequently.[9] Interestingly, Davis (2006) finds that larger population within 10 miles increases demand to a movie theater, and that the geographical market of a theater extends for at most 15 miles around it: we find for the same industry that 75% of the transactions occur in fact within (around) 11 miles.

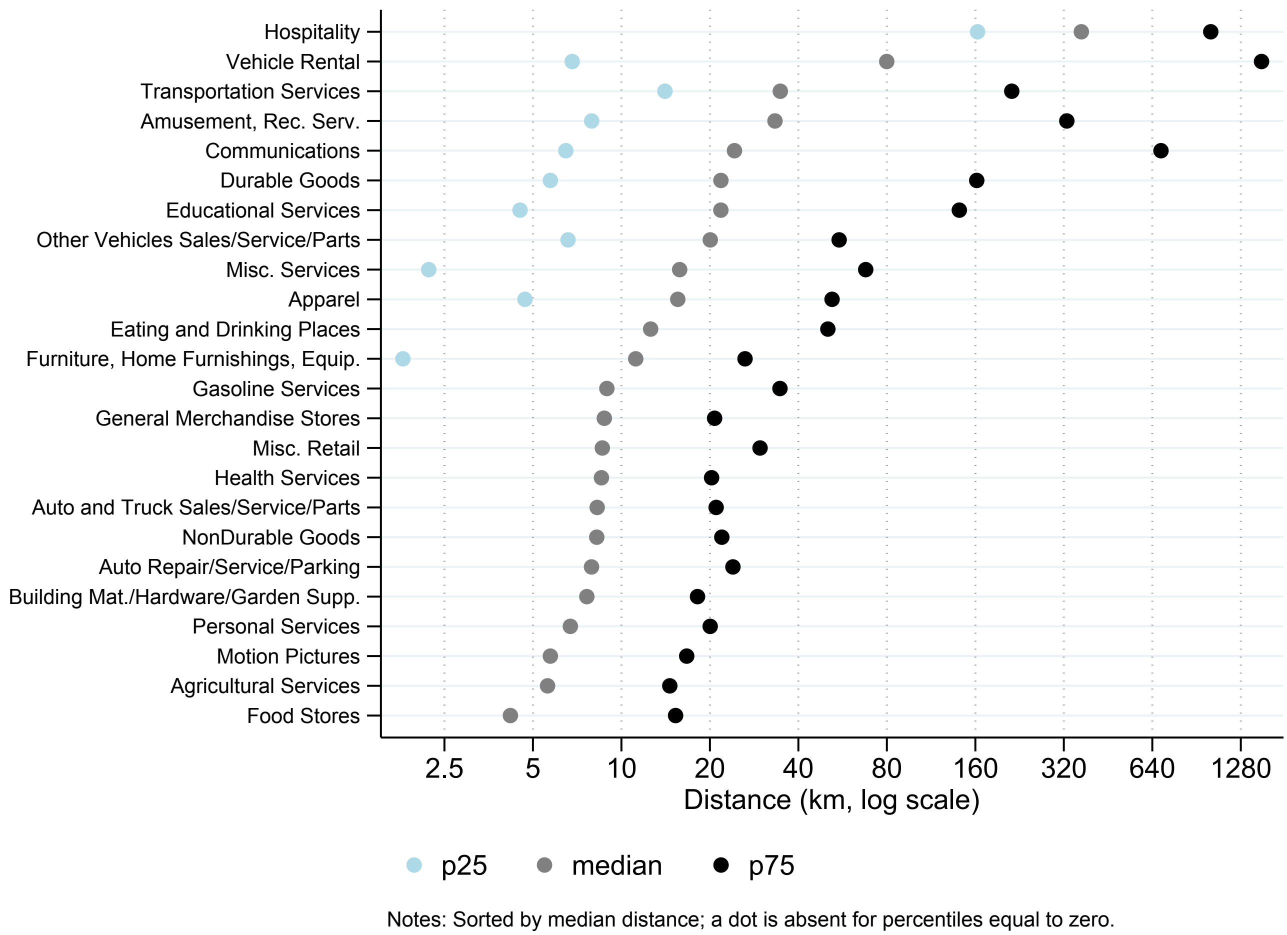

Figure 2: Distances traveled by sector (select percentiles)

Obviously, the measured distance traveled by consumers is a combination of their willingness to travel (as mediated by their optimal shopping behavior) and supply conditions like the density of producers. We will return to this distinction later. For now, we emphasize that the spatial dimension of consumers' behavior is actively moving in the data: consumers visit just a few locations among the many available, but the typical distance traveled varies broadly across sectors. To understand more the local nature of

[9] We will show below a more precise relation between the importance of distance and the frequency of transactions.

different consumption markets, we need to explore further the determinants of the relation between total purchases and distance. We move to this task next.

3.4 Gravity in consumer expenditure

Gravity is an almost universal feature of spatial relationships.[10] While a large literature has documented the decay of goods' trade flows with distance at inter-national and intra-national level, little is known about the spatial behavior of consumers.[11] We fill this gap in two steps.

First, we document that gravity also holds for consumers' behavior. We make full use of the information available in the data comparing 1) expenditure inside vs. outside one's place of residence, and 2) the decline in expenditure across merchants at different distances from home.

We then analyze the determinants of this decline, decomposing the total decay into fewer accounts transacting (an extensive "accounts" margin) and lower average expenditure per account (an intensive "expenditure" margin). The decline in average expenditure per account may be further decomposed into average expenditure per transaction (a "batch size" margin) and number of transactions (a "frequency" margin).

3.4.1 Expenditure patterns display gravity

We start our exploration of gravity by investigating how quickly total expenditure decays with distance. A large empirical literature has documented that merchandise trade flows decay with distance both across countries (e.g., Disdier and Head 2008) and within countries (e.g., Hillberry and Hummels 2007). Since final consumers buy goods directly from producers only in a minority of cases, our knowledge of gravity in final consumption is extremely limited. Moreover, virtually all the literature deals with merchandise shipments, thus ignoring the service sector altogether. Here, we fill these important gaps.

Denote with $x_{a,m}$ the observed expenditure of account a falling on merchant m in a sector in the whole sample period. Note that each account a has an associated home location $h = r(a)$, and each merchant has a sale location $s = l(m)$. We start by aggregating our expenditure at the sector – account home – sales location level, X_{hs}.

$$X_{hs} \equiv \sum_{a:r(a)=h,m:l(m)=s} x_{a,m}$$

We initially relate the expenditure X_{hs} to the distance between residence and merchant's shop in two ways. First, we simply estimate the change in expenditure associated with shopping out of the home residence:

$$\log X_{hs} = \alpha + \gamma^{(h)} + \gamma^{(s)} + \eta \times \mathbf{1}_{(h \neq s)} + \varepsilon_{hs} \tag{18}$$

[10] Disdier and Head (2008) argue that in the international trade literature, a typical elasticity of trade flows to distance is around −1. Within country estimates are slightly higher: Hillberry and Hummels (2007) estimate a coefficient of around −1.3 on short distances with microdata on shipment of firms. Monte, Redding and Rossi-Hanbserg (2016) find an elasticity of commuting flows to distance of −4.4.

[11] International flows of goods are only measured at country level, thus ignoring the travel dimension of consumers' purchase. Intra-national flows of goods typically record firm-to-firm transactions.

where $\mathbf{1}_{(h \neq s)}$ is an indicator function assuming the value of 1 if $h \neq s$ and zero otherwise. Second, we follow the gravity literature and estimate the impact of distance on trade flows with a regression of the form

$$\log X_{hs} = \alpha + \gamma^{(h)} + \gamma^{(s)} + \delta \log dist_{hs} + \varepsilon_{hs} \tag{19}$$

including only pairs where $h \neq s$; in this equation, α is a constant, and $dist_{hs}$ is the distance between the centroids of h and s. In both equations, a set of origin and destination fixed effects, $\gamma^{(h)}$ and $\gamma^{(s)}$, controls for unobserved differences in productivity and intensity of competition (Anderson and Van Wincoop, 2003). These two approaches highlight complementary features of the data: eq. (18) shows the importance of very short trips, for which we don't have good measures of distances; eq. (19) shows the elasticity of expenditure to distance, when distance can be measured.

We first estimate equations (18) and (19) across all sectors, using distances up to 120 km. We find, unsurprisingly, very clear effects of distances. Estimating (18), the average expenditure out of home is about 8.8% of the average expenditure at home ($\eta = -2.428$, robust s.e. 0.021).[12] When we estimate (19), we find a slope of -1.049 (s.e. of 0.006), very much in line with estimates in the trade literature.[13]

These estimates mask large differences across sectors. Table 4 shows the coefficients of η (column 1) and δ (column 4) when we estimate eq. (18) and (19) by sector. Sectors in this table are ordered by the out-of-home dummy in column 1 (this ordering will be kept throughout the paper for ease of reference). A large fraction of the total decay appears to occur on short distances already, as implied by fairly high estimates of η. However, such decay is heterogeneous: in sectors like Food Stores, the point estimate of average expenditure out of home is around 10% the expenditure at home; this fraction grows to 20% for eating and drinking places, 38% for personal services, and at 91% for Durable Goods. The impact of distance as measured by estimates of eq. (19) is consistent with this picture: the correlation between the two sets of coefficients is 0.68.

3.4.2 Margins of Adjustment

Why does expenditure decay with space? Our simple models points to two main sources: fewer people traveling out, and lower expenditure per person; furthermore, expenditure per person can decrease because of smaller batch sizes, or less frequent trips. These sources map well into simple decompositions in the spirit of Hummels and Klenow (2005) and Hillberry and Hummels (2007). In any given sector, we express total expenditure of consumers in h falling on merchants in s as

$$X_{hs} = \underbrace{N_{hs}}_{\text{account margin}} \times \underbrace{\bar{x}_{hs}}_{\text{expenditure margin}} = \tag{20}$$

$$= \underbrace{N_{hs}}_{\text{account margin}} \times \underbrace{f_{hs}}_{\text{frequency margin}} \times \underbrace{\bar{x}_{hs}/f_{hs}}_{\text{batch size margin}} \tag{21}$$

[12] Using all data, we find $\eta = -2.545$ (robust s.e. 0.0223).

[13] This slope is not particularly sensitive to changes in the cutoff. See Appendix C.4, page 35, for further discussion.

Table 4: **Decline in expenditure**

Category	Out of Home coeff (1)	Out of Home pv (2)	Out of Home obs. (3)	Gravity coeff (4)	Gravity pv (5)	Gravity obs. (6)
Food Stores	-2.23	0.00	22,652	-0.85	0.00	18,635
Gasoline Services	-2.08	0.00	39,673	-0.60	0.00	34,621
General Merchandise Stores	-1.79	0.00	26,845	-0.93	0.00	23,933
Misc. Retail	-1.70	0.00	34,057	-0.65	0.00	30,046
Eating and Drinking Places	-1.57	0.00	34,509	-0.56	0.00	31,028
Building Mat./Hardware/Garden Supp.	-1.40	0.00	14,190	-0.73	0.00	11,610
Auto Repair/Service/Parking	-1.25	0.00	4,415	-0.40	0.00	3,014
NonDurable Goods	-1.16	0.00	978	-0.65	0.00	758
Health Services	-1.12	0.00	5,136	-0.33	0.00	3,914
Apparel	-1.10	0.00	15,921	-0.53	0.00	14,069
Transportation Services	-1.09	0.00	743	-0.47	0.00	635
Furniture, Home Furnishings, Equip.	-1.07	0.00	12,292	-0.57	0.00	10,740
Auto and Truck Sales/Service/Parts	-1.04	0.00	7,302	-0.33	0.00	5,508
Motion Pictures	-1.04	0.00	1,927	-0.34	0.00	1,253
Amusement, Rec. Serv.	-1.02	0.00	2,959	-0.22	0.00	2,330
Educational Services	-1.00	0.00	712	-0.15	0.38	530
Personal Services	-0.96	0.00	5,204	-0.31	0.00	3,761
Vehicle Rental	-0.95	0.00	546	-0.08	0.59	296
Misc. Services	-0.92	0.06	222	0.97	0.01	120
Communications	-0.89	0.00	424	-0.41	0.01	263
Agricultural Services	-0.88	0.00	552	0.42	0.11	190
Other Vehicles Sales/Service/Parts	-0.68	0.41	257	-0.59	0.08	128
Hospitality	-0.65	0.01	1,394	-0.14	0.08	1,160
Durable Goods	-0.09	0.90	79	1.11	0.67	15

Eq. (20) says that as distance increases, expenditure can decrease either because the number of agents traveling decreases (the extensive "account margin") or because agents spend less on average (the intensive "expenditure margin"). In turn, lower expenditure per account on average can arise either because each transaction is smaller (the "batch size margin") or because consumers transact less often (the "frequency margin"), as emphasized in (21). When we re-estimate eq. (18) with the left hand-side being each of these three terms, the coefficients on the out-of-home dummy add up to the overall coefficients η reported in column 1 of Table 4 (and similarly for eq. (19)).[14]

Figure 3 shows the results of this decomposition for eq. (18). The length of each bar corresponds to column 1 in Table 4. The blue bar measures the contribution of the account margin. For the typical

[14] A further angle of this decomposition could relate to the Allen and Alchian (1964) conjecture: consumers should be willing to travel more for higher quality goods and services when travel costs do not vary with quality. Hence, there should be a positive relation between average value of a transaction and distance. Unfortunately, our data does not allow a precise mesurement of unit values and hence cannot be used to speak to this conjecture. For related work, see Hummels and Skiba (2004).

sector, 72% of the drop in out-of-home expenditure is associated with fewer people traveling outside, rather than to people spending less on average for out-of-home transactions.[15] Through the lenses of our simple model, the account margin drop is related to the extent of heterogeneity in travel costs around marginal consumers indifferent between purchasing at home or traveling outside (i.e., the density $T'(t_c)$): the response of this margin is stronger when there are many consumers with similar travel costs around t_c. As a benchmark, Hillberry and Hummels (2007) find, for firm-to-firm shipments within U.S., that on short distances the extensive margin explains almost the totality of the decay. We take this as evidence that our "expenditure" margin is active in consumers in a way that it is not for producer-to-producer transactions: while variation in trade costs impacts expenditure per firm within U.S. minimally, variation in travel costs impacts expenditure per person substantially.

Again, however, we find significant differences across sectors. For Personal Services or Motion Pictures, for example, almost the totality of the fall is due to fewer people traveling outside; For Food Stores, on the other hand, fewer people traveling outside only explains half of the fall in expenditure occurring out-of-home. What accounts for the rest of the drop?

Figure 3 indicates that across all sectors, average expenditure per account (the expenditure margin) drops outside of home almost exclusively because of the frequency margin: consumers choose to travel outside less frequently. The drop in the average transaction value (the batch size margin) has a limited role in most cases. Tables C.7 and C.8 in the Appendix (p. 41 and 41) show that the combination of the account and frequency margin typically contribute 90-95% of the decline in expenditure.[16]

What can explain this heterogeneity in gravity? Our model points to one source of sector heterogeneity, differences in storage costs, that links the strength of gravity to the sample average frequency of transactions. To take a closer look at this conjecture, we plot the out-of-home expenditure as a share of home expenditure $\exp(\eta)$ (using column 1 in Table 4) against the average number of transactions per account in the sector from the data. Figure 4 shows this relation for the sectors where the out-of-home dummy is statistically significant from zero.[17] When customers (optimally) choose more visits, gravity is more important. Note that since the average number of transactions has not been used directly to compute the out-of-home dummy, there is nothing mechanical about this empirical relation.

Our simple model provides a possible explanation for this correlation. Buying outside of the home place requires, *ceteris paribus*, larger batch sizes: hence, if storage costs in a sector are higher, consumers want to reduce average inventory. They do so in two ways: they buy more of the overpriced varieties at home and fewer varieties outside (the expenditure margin), or stop traveling outside altogether (the account margin). This behavior maps into stronger gravity (Proposition 2). On the other hand, however, lower inventory requires more trips (through Proposition 3). This response generates a negative correlation

[15] Figure C.2 in the Appendix (p. 42) shows the same decomposition for eq. (19): the extensive margin is even less important, accounting for 57% of the decline in expenditure in the typical sector as distance increases. Tables C.5 and C.6, also in the Appendix (pages 40 and 40), show the actual values of the account and expenditure margin with associated p-values.

[16] This feature of the data explains the emphasis, in our model, of storage costs and endogenous frequency of transactions.

[17] Figures C.3, C.4 and C.5 in the Appendix, starting at page 43, replicates it for all the sectors and for the impact of distance using gravity.

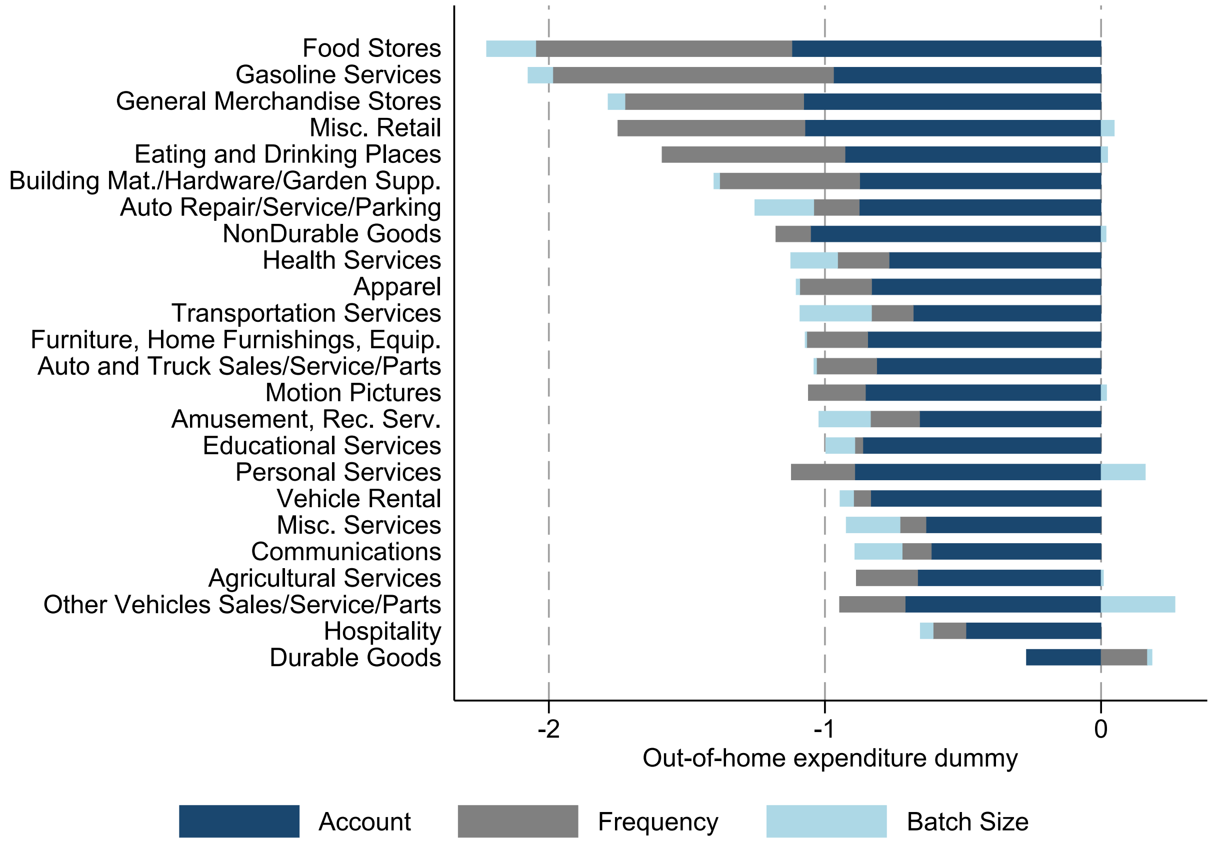

Figure 3: Margins in the out-of-home expenditure drop

between the strength of gravity and frequency of transactions, as present in Figure 4.

These results provide the first piece of (somewhat indirect) evidence that demand conditions may matter in determining local equilibrium outcomes. In sectors where storage costs are high, consumer demand declines faster with distance: from a merchant's point of view, the market will be more localized, and distant competition will be less of a threat.

In the next section we consider a more direct shock to travel cost.

4 The Effect of Rain

The analysis up to now has shown that consumers' typical travel ranges are limited, expenditure declines with distance, and the combination of accounts and frequency margins are the main reasons of such decline. We have also shown that the strength of gravity varies by sector, and in a way that can be related to sector-level characteristics like storage costs and durability.

While certainly suggestive, these facts per se do not show that consumers' willingness to travel has a role in determining local product market outcomes. Some sector-level characteristics of the supply side

Figure 4: Drop in expenditure out of home

may bring producers and consumers closer to each other, so that the observed distance traveled is shorter. One such characteristic could be the fixed costs of operating a store: everything else equal, high fixed costs would imply fewer suppliers, so that consumers would need to travel more on average.

To make progress on this issue, we need a plausible shifter to consumers' travel costs whose variation is uncorrelated to residential decisions of consumers and locations decisions of firms in the sample period. We can then study the impact of this cost shifter on the spatial distribution of observed expenditure.

We turn to rain. We use daily data on rainfall precipitation from the National Oceanic and Atmospheric Administration, as described in Menne et al. (2012). For each centroid of a residence or a shopping location in our data, we find the closest weather station among the roughly twelve thousand disseminated over the U.S. territory. In the transaction data, the median distance between a weather station and a merchant is 6.5 km (mean 7.3 km) and the median distance to a residence is 7.3 km (mean 8 km).

We use daily data on rainfall to assign a weather status for each transaction. We create a transaction-level indicator variable that assumes the value of 1 if, during the transaction day, the associated weather stations recorded rain both in the residence and in the shopping location. During the sample period, 28% of transactions have a rain episode so defined. A concern could be that most of the variation in this indicator is geographically related, rather than occurring within residence–location pairs over time. This is not the case. A regression of the weather status indicator variable on residence − shopping location pairs and transaction date fixed effects absorbs 24% of the variation in the transaction level data, leaving ample residual variation to identify movements in the spatial distribution of expenditure.

We recompute for each pair of locations h,s two expenditure flows: one observed during non-rainy days, and one observed during rainy days. At this point, we are in a position to extend our analysis from equations (18) and (19) above to include the effect of rain. Specifically, we start by estimating:

$$\log X_{hs} = \alpha + \gamma^{(h)} + \gamma^{(s)} + \eta \times \mathbf{1}_{(h \neq s)} + \rho \times \mathbf{1}_{RAIN} + \mu \times \left(\mathbf{1}_{(h \neq s)} \times \mathbf{1}_{RAIN}\right) + \varepsilon_{hs} \tag{22}$$

where $\mathbf{1}_{RAIN}$ is an indicator variable assuming the value of 1 if the observation refers to rainy days and zero otherwise. The presence of origin and destination fixed effects ensures that average levels of rain by location are not contributing to the identification.[18] Table 5 shows the estimated values of ρ (column 1), η (column 3) and μ (column 5).

Rain affects expenditure at home significantly for most sectors. In the median sector, expenditure at home on a rainy day is about $100 \cdot \exp(-0.47) = 63\%$ of the expenditure on a non–rainy day. Food Stores, Building Materials and Garden Supplies, and General Merchandise are the most impacted; Communications, Transportation Services are the least impacted. Note that the ratio of ρ between a sector at the 10% of impact and one at the 90% is around 1.7 (and the response of expenditure between the least and the most impacted sector varies by a factor of 5): a common cost shock to all sectors induces differential responses across sectors.

Rain also implies a drop in expenditure out of home: $\rho + \mu$ is negative for all sectors. In the typical sector, expenditure outside on rainy days is 77% of the expenditure outside on non–rainy days. The heterogeneity in the responses stays in the same order of magnitude: the $p90/p10$ ratio in $\rho + \mu$ is around 1.6, and the max/min ratio is around 4.5.

Not only does a cost shock impact the levels of consumers expenditure, but it does so differentially over space. Column 5 reveals that rain impacts the spatial distribution of expenditure making the decline in expenditure out of home *less* pronounced. This behavior is consistent with the shock selecting the type of agents going out for shopping. While some agents do not shop (the expenditure declines both at home and outside), those who choose to go out must be those less sensitive to travel costs, and hence the composition of expenditure shifts toward out-of-home locations. For the median sector, the out-of-home expenditure decline in rainy conditions is 19% flatter than in non-rainy days.

It is useful to stress an implication of these results. These regressions are showing that selection along travel costs is an important determinant of the spatial distribution of expenditure given location decisions of consumers and firms. In other words, the level and the elasticity of demand faced by merchants in a

[18] In this residence location – shopping location – weather status dataset, a regression of the weather status indicator on residence – shopping location pair fixed effects has an R^2 of 20%. The effect of rain can then be identified comparing the same pair of locations in rainy and non–rainy days.

Table 5: **Expenditure out of home place and rain**

Category	Rain coeff	Rain pv	Out of home coeff	Out of home pv	Out of home × Rain coeff	Out of home × Rain pv	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Food Stores	-0.60	0.00	-2.02	0.00	0.29	0.00	35,942
Gasoline Services	-0.54	0.00	-1.84	0.00	0.29	0.00	55,990
General Merchandise Stores	-0.60	0.00	-1.57	0.00	0.29	0.00	39,601
Misc. Retail	-0.55	0.00	-1.51	0.00	0.26	0.00	48,146
Eating and Drinking Places	-0.47	0.00	-1.36	0.00	0.20	0.00	46,545
Building Mat./Hardware/Garden Supp.	-0.65	0.00	-1.20	0.00	0.25	0.00	22,298
Auto Repair/Service/Parking	-0.50	0.00	-1.12	0.00	0.27	0.00	7,296
NonDurable Goods	-0.38	0.00	-0.86	0.00	0.09	0.45	2,277
Health Services	-0.51	0.00	-0.99	0.00	0.17	0.02	8,412
Apparel	-0.44	0.00	-0.94	0.00	0.20	0.00	21,857
Transportation Services	-0.14	0.43	-0.75	0.01	-0.12	0.52	1,493
Furniture, Home Furnishings, Equip.	-0.52	0.00	-0.90	0.00	0.19	0.01	17,445
Auto and Truck Sales/Service/Parts	-0.49	0.00	-0.83	0.00	0.25	0.00	11,519
Motion Pictures	-0.43	0.00	-0.90	0.00	0.19	0.00	3,861
Amusement, Rec. Serv.	-0.35	0.00	-0.88	0.00	0.11	0.26	4,470
Educational Services	-0.47	0.00	-0.90	0.00	0.22	0.22	1,564
Personal Services	-0.42	0.00	-0.76	0.00	0.12	0.02	8,880
Vehicle Rental	-0.44	0.00	-0.83	0.00	0.36	0.02	1,310
Misc. Services	-0.41	0.01	-0.91	0.01	0.16	0.39	1,044
Communications	-0.12	0.40	-0.57	0.00	-0.11	0.51	1,005
Agricultural Services	-0.44	0.00	-0.75	0.00	0.13	0.14	2,517
Other Vehicles Sales/Service/Parts	-0.58	0.00	-0.60	0.29	0.44	0.03	1,082
Hospitality	-0.46	0.01	-0.53	0.01	0.24	0.21	2,219
Durable Goods	-0.52	0.01	-0.14	0.72	0.15	0.53	592

particular location will depend, among other things, on the level and the heterogeneity in travel costs of consumers living in proximity to such location: to the extent that rational firms are aware of these costs, local demand conditions must matter for local equilibrium outcomes.

We also find that the spatial distribution of expenditure moves differentially across sectors. A simple statistic to consider would compare $\exp(\eta)$, the percentage drop in out-of-home expenditure vs. home expenditure without rain, to $\exp(\eta+\mu)$, giving the same percentage in rainy days. In Figure 5 we plot $\exp(\eta+\mu)-\exp(\eta)$ against the frequency of transactions.[19]

We find large differences in the impact of rain across sectors. A common travel cost shock makes gravity flatter by 4.4 percentage points in Food Stores, 8.7 points in Motion Pictures, 14 points in Durable Goods and 18.6 points in Vehicle rentals. The spatial distribution of demand is impacted less for sectors with more frequent purchases. This finding provides some validation to Figure 4 above. When a good is less

[19] This figure excludes one outlier, Other Vehicle Sales, Services and Parts, which is imprecisely estimated. This sector includes items like Boat, Motorcycles, or Camper Dealers, for example. For this sector, the estimated difference is 30 percentage points. Figure C.6 in Appendix, p. 44, shows the complete picture.

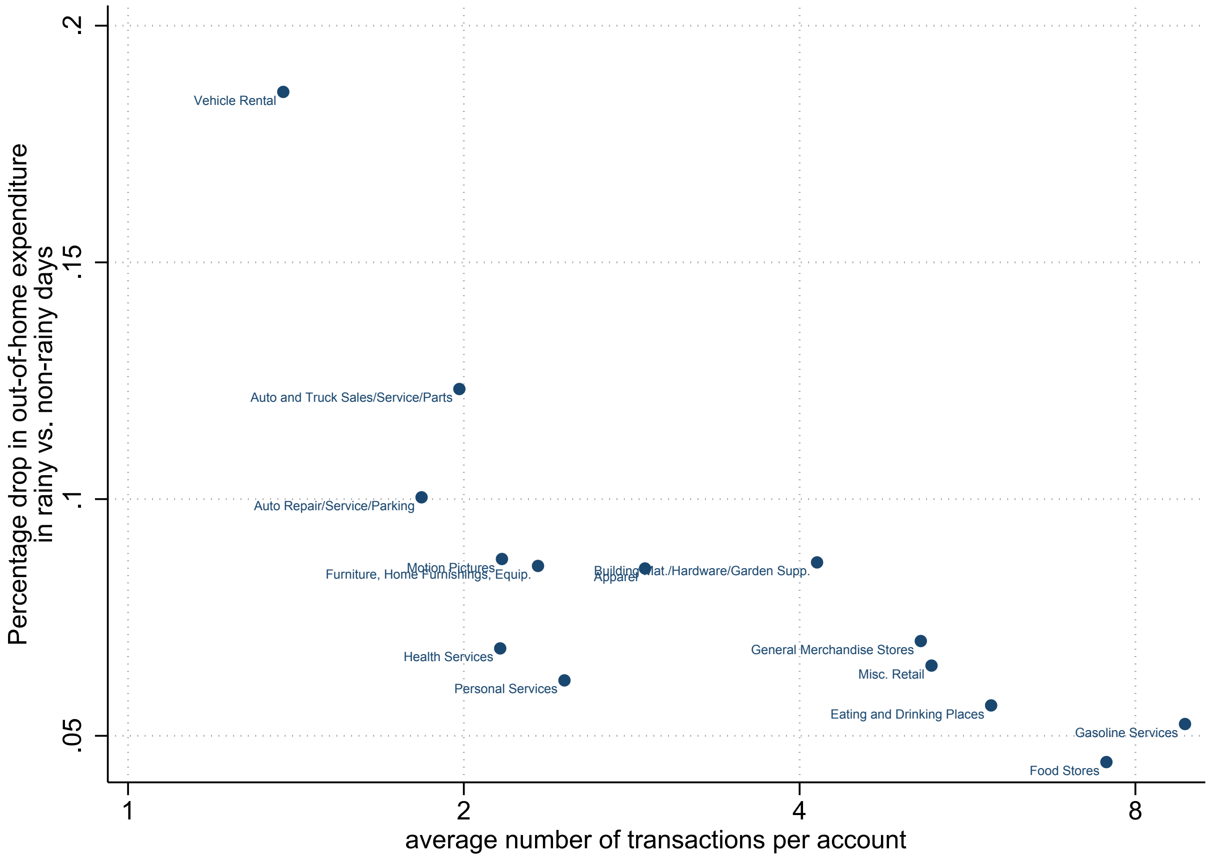

Figure 5: The flattening in gravity across sectors

storable, purchases are more frequent, less sensitive to rain episodes, but also more local: gravity matters more.

A final insight into the impact of rain can be gauged by comparing Table 5 with the results of the following gravity regression:

$$\log X_{hs} = \alpha + \gamma^{(h)} + \gamma^{(s)} + \delta \log dist_{hs} + \tilde{\rho} \times \mathbf{1}_{RAIN} + \tilde{\mu} \times (\mathbf{1}_{RAIN} \times \log dist_{hs}) + \varepsilon_{hs} \qquad (23)$$

This regression augments eq. (19) with the weather status indicator, in levels and interacting it with distance. It estimates the effect of rain by comparing only places out of home at different distances. Table 6 reports the results. The most notable difference with respect to Table 5 above is that the interaction between distance and rain is now very small and almost everywhere statistically insignificantly different from zero. Reasonably, rain alters the spatial composition of expenditure, but only by leaving at home those whom would take short trips *more* than those willing to travel longer: once expenditure is no longer occurring at very short distances, an extra kilometer of travel does not matter.

Table 6: **Gravity and rain**

Category	Rain		Distance		Distance × Rain		Obs.
	coeff	pv	coeff	pv	coeff	pv	
Food Stores	-0.40	0.00	-0.71	0.00	-0.00	0.98	28,485
Gasoline Services	-0.43	0.00	-0.50	0.00	0.03	0.00	47,006
General Merchandise Stores	-0.59	0.00	-0.77	0.00	0.07	0.00	34,539
Misc. Retail	-0.39	0.00	-0.50	0.00	0.01	0.45	41,449
Eating and Drinking Places	-0.44	0.00	-0.45	0.00	0.04	0.00	40,855
Building Mat./Hardware/Garden Supp.	-0.48	0.00	-0.56	0.00	0.01	0.69	17,873
Auto Repair/Service/Parking	-0.00	0.99	-0.28	0.00	-0.09	0.07	4,733
NonDurable Goods	-0.23	0.25	-0.48	0.00	-0.03	0.67	1,659
Health Services	-0.19	0.13	-0.18	0.00	-0.06	0.20	6,234
Apparel	-0.37	0.00	-0.40	0.00	0.03	0.13	18,979
Transportation Services	-0.23	0.35	-0.34	0.00	-0.02	0.85	1,251
Furniture, Home Furnishings, Equip.	-0.49	0.00	-0.41	0.00	0.05	0.10	14,964
Auto and Truck Sales/Service/Parts	-0.30	0.02	-0.20	0.00	0.01	0.80	8,290
Motion Pictures	-0.16	0.22	-0.21	0.00	-0.05	0.36	2,393
Amusement, Rec. Serv.	-0.38	0.01	-0.18	0.00	0.04	0.40	3,399
Educational Services	-0.44	0.15	-0.15	0.31	0.06	0.54	1,106
Personal Services	-0.21	0.02	-0.20	0.00	-0.04	0.25	6,299
Vehicle Rental	-0.01	0.98	-0.03	0.83	-0.04	0.70	741
Misc. Services	0.11	0.75	1.16	0.00	-0.15	0.19	602
Communications	0.35	0.28	-0.27	0.04	-0.22	0.04	634
Agricultural Services	-0.47	0.02	0.64	0.00	0.06	0.45	1,350
Other Vehicles Sales/Service/Parts	-0.15	0.69	-0.45	0.09	-0.01	0.97	702
Hospitality	-0.17	0.55	-0.11	0.12	-0.02	0.82	1,855
Durable Goods	-0.35	0.40	1.70	0.30	-0.01	0.94	365

5 Conclusion

Using detailed geographical information from more than 1.7 million individual consumers' credit card transactions, we document several stylized facts regarding the geography of consumption. We find large heterogeneity across industries in the overall impact of distance and in the importance of extensive margins. We also find that the differences in gravity across industries are correlated with the frequency of transactions. A simple model of consumer choice suggests that this correlation can be induced by heterogeneity in the durability/storability of the final item demanded, which we interpret as a new determinant of gravity. We argue that storable/durable final consumption items may have an inherently larger geographical markets than perishable/non-durable goods. Hence, the distribution of travel costs of consumers will matter differentially for firms operating in different sectors.

Since unobserved heterogeneity in sector characteristics might be a contributing factor in determining gravity patterns, we also provide more direct evidence on the response of the spatial distribution of expenditure to travel costs shocks. We find that rain reduces expenditure of consumers living in a particular place both at home and in outside locations, but differentially over space and across sectors.

In particular, gravity becomes flatter, since only consumers less sensitive to travel costs decide to go out and spend; moreover, rains matters more for more durable/storable goods. Again, the joint distribution of consumers and their travel costs matters for the level and elasticity of demand faced by any merchant.

Taken together, these results indicate that local demand conditions are quantitatively relevant for local equilibrium outcomes and for their response to local shocks. Incorporating demand-side characteristics is essential to analyzing firms' location and production decisions, in particular in the understudied service sector, which accounts a large share of economic activity. Further, our results provide important information for the study of the liberalization of international trade and investment in services: entry and location decisions of foreign firm establishments in a local market will be shaped, among other things, by the different degrees of localization of their product's market.

References

[1] Alchian, A., and W. Allen (1964). "University Economics". Belmont, Calif.: Wadsworth.

[2] Agarwal, S., N. Marwell, and L. McGranahan (forthcoming). "Consumption Responses to Temporary Tax Incentives: Evidence from State Sales Tax Holidays", American Economic Journal: Economic Policy.

[3] Anderson, J. (1979). "A Theoretical Foundation for the Gravity Equation", American Economic Review, 69:1, 106:116.

[4] Anderson, J. and E. Van Wincoop (2003). "Gravity with gravitas: a solution to the border puzzle", American Economic Review, 93:1, pp. 170-192

[5] Bernard, A., B. Jensen, S. Redding and P. Schott (2010). "Wholesalers and Retailers in U.S. Trade (Long Version)", NBER Working Paper 15660.

[6] Chandra, A., K. Head and M. Tappata (2014). "The Economics of Cross-Border Travel", The Review of Economics and Statistics 96:4 2014, pp. 648-661.

[7] Couture, V. (2016). "Valuing the Consumption Benefits of Urban Density", Working paper.

[8] Disdier, A. and K. Head (2008). "The puzzling persistence of the distance effect on bilateral trade", Review of Economics and Statistics, 90:1, pp. 37-48.

[9] Davis, D., J. Dingel, J. Monras, and E. Morales (2016). "How Segregated is Urban Consumption?", Working paper.

[10] Davis, P. (2006). "Spatial Competition in Retail Markets: Movie Theaters", RAND Journal of Economics, 37:4, pp. 964-982.

[11] Dornbush, R., S. Fischer and P. Samuelson (1977). "Comparative advantage, trade, and payments in a ricardian model with a continuum of goods", American Economic Review, 67:5, pp. 823-839.

[12] Eaton, J., and S. Kortum (2003).. "Technology, Geography, and Trade", Econometrica, 70:5, pp. 1741-1779.

[13] Ellison, G. and S. F. Ellison (2009). "Tax Sensitivity and Home State Preferences in Internet Purchasing", American Economic Journal: Economic Policy, 1:2, p. 53-71.

[14] Glaeser, E., J. Kolko, and A. Saiz (2001). "Consumer City", Journal of Economic Geography, Vol. 1, pp. 27-50.

[15] Handbury, J., I. Rahkovsky, and M. Schnell (2016). "Is the Focus on Food Deserts Fruitless? Retail Access and Food Purchases Across the Socioeconomic Spectrum", Working paper.

[16] Hillberry, R. and D. Hummels (2007). "Trade responses to geographic frictions: A decomposition using micro-data", European Economic Review 52:3, pp. 527:550

[17] Hummels, D. and P. Klenow (2005). "The variety and quality of a nation's exports", American Economic Review, 95:3, pp. 704-723.

[18] Hummels, D. and A. Skiba (2004). "Shipping the good apples out? An empirical confirmation of the Alchian-Allen conjecture", Journal of Political Economy, 112:6, pp. 1384-1402.

[19] Hortaçsu, A., F. A. Martínez-Jerez, and J. Douglas (2009). "The Geography of Trade in Online Transactions: Evidence from eBay and MercadoLibre", American Economic Journal: Microeconomics, 1:1, p. 53-74.

[20] Hortaçsu, A. and C. Syverson (2015). "The ongoing evolution of US retail: A format tug-of-war", Journal of Economic Perspectives, 29:4, pp. 89-112.

[21] Houde, J.F. (2012). "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline", American Economic Review, 102:5, pp. 2147-2182.

[22] Jarmin, R., S. Klimek, and J. Miranda (2005). "The Role of Retail Chains: National, Regional and Industry Results", in "Producer Dynamics: New Evidence from Micro Data", University of Chicago Press, p. 237-262.

[23] Menne, M.J., I. Durre, R.S. Vose, B.E. Gleason, and T.G. Houston (2012). "An overview of the Global Historical Climatology Network-Daily Database", Journal of Atmospheric and Oceanic Technology, 29, pp. 897-910.

[24] Monte, F., S. Redding and E. Rossi-Hansberg (2016). "Commuting, migration, and local employment elasticities", Working paper.

[25] Nakamura, E. (2008). "Pass-Through in Retail and Wholesale", American Economic Review: Papers and Proceedings, 98:2, 430-437.

[26] Oi, W. (1992). "Productivity in the distributive trades: the shopper and the economies of massed reserves", in "Output measurement in the services sectors", University of Chicago Press, p. 161-193.

[27] Reilly, W. (1931). "The Law of Retail Gravitation", New York: Knickerbocker Press.

[28] Schiff, N. (2015). "Cities and Product Variety: Evidence from Restaurants", Journal of Economic Geography, Vol 15, pp. 1085-1123.

A Analytical Appendix

Proposition 1. *If $d(j)$ is strictly concave for $j \geq 1/2$, there exist a unique level of differences in travel costs t_c such that a consumer ω travels if and only if $t(\omega) \leq t_c$.*

Proof. Note from (13) that the optimal j^* can be expressed as

$$j^*(t) : \min\left\{d(j^*) = \left(\frac{\bar{q}g}{2}\right)^{1/2} t, 1\right\} \tag{A.1}$$

where we have defined

$$t \equiv \left[t_s(\omega)^{1/2} - t_h^{1/2}\right] \tag{A.2}$$

If $t = 0$ (that is, for consumers where the home and outside cost of travel are identical), $d(j^*) = 0 \Longrightarrow j^* = j_z = 1/2$. When $t \geq t_{\max} \equiv \left(\frac{2}{\bar{q}g}\right)^{1/2} d(1)$, then $j^*(t) = 1$ always: for high enough travel costs, consumers will always choose not to travel and buy everything at home. Note that

$$\frac{dj^*}{dt} = \left(\frac{\bar{q}g}{2}\right)^{1/2} \frac{1}{d'(j^*)} > 0 \text{ for } t \leq t_{\max} \tag{A.3}$$

and

$$\frac{d}{dt}\frac{dj^*}{dt} = -\left(\frac{\bar{q}g}{2}\right)^{1/2} \frac{1}{d'(j^*)^2} d''(j^*) \frac{dj^*}{dt} = -\left(\frac{\bar{q}g}{2}\right) \frac{d''(j^*)}{d'(j^*)^3}$$

Hence, if $d(j)$ is strictly concave, $j^*(t)$ is strictly convex in t.

The cutoff condition in (15) above states that a consumer will be willing to travel if

$$\gamma (g\bar{q})^{1/2} t (1 - j^*(t)) \leq \int_{j^*(t)}^{1} d(j)\, dj \tag{A.4}$$

where we expressed j^* as a function of t and used (A.2). We define t_c as the travel cost difference satisfying this condition (A.4) with equality. At $t = 0$, the right-hand side starts strictly above zero, it is strictly decreasing for $t \leq t_{\max}$ (since it represents the total savings on all varieties purchased outside), it reaches 0 at $t = t_{\max}$, and stays constant at 0 after that.

Define the left-hand side function as

$$l(t) \equiv \gamma (g\bar{q})^{1/2} t (1 - j^*(t))$$

This function is such that

$$\begin{aligned} l'(t) &= \gamma (g\bar{q})^{1/2} \left[1 - j^*(t) - t j^{*\prime}(t)\right] \\ l''(t) &= \gamma (g\bar{q})^{1/2} \left[-j^{*\prime}(t) - j^{*\prime}(t) - t j^{*\prime\prime}(t)\right] \end{aligned}$$

Hence, the left-hand side of (A.4) is such that $l(0) = l(t_{\max}) = 0$, $l'(0) > 0$ and $l'(t_{\max}) < 0$. From above, if $d(j)$ is strictly concave then j^* is strictly convex, and hence $l''(t) < 0\ \forall t \in [0, t_{\max}]$; that is, if $d(j)$ is strictly concave then $l(t)$ is also strictly concave over the range $[0, t_{\max}]$. This implies that the inequality (A.4) is satisfied for a range $t \in [0, t_c]$, where t_c is the unique intersection of the left-hand side and the right-hand side. ■

Proposition 2. *If storage costs increase, the travel cutoff t_c decreases, i.e., $dt_c/dg < 0$.*
Proof. From (13), we have that

$$\frac{dj^*}{dg} = \frac{1}{2}\left(\frac{\bar{q}}{2g}\right)^{1/2} \frac{t}{d'(j^*)} \tag{A.5}$$

as long as $t \leq t_{\max}$. Recall that the definition of the cutoff t_c is

$$t_c : \gamma g^{1/2} \bar{q}^{1/2} t_c = \frac{\int_{j^*(t_c)}^{1} d(j)\, dj}{1 - j^*(t_c)} \equiv rhs\left(j^*(t_c)\right) \tag{A.6}$$

Totally differentiating this expression with respect to t_c and g,

$$\gamma g^{1/2} \bar{q}^{1/2} dt_c + \frac{1}{2}\gamma \left(\frac{\bar{q}}{g}\right)^{1/2} t_c dg = rhs' \frac{dj^*}{dt_c} dt_c + rhs' \frac{dj^*}{dg} dg \Longrightarrow$$

$$\frac{dt_c}{dg} = \frac{rhs' \frac{dj^*}{dg} - \frac{1}{2}\gamma\left(\frac{\bar{q}}{g}\right)^{1/2} t_c}{\gamma g^{1/2}\bar{q}^{1/2} - rhs' \frac{dj^*}{dt_c}}$$

Using the expressions above (A.3) and (A.5),

$$\frac{dt_c}{dg} = \frac{rhs' \frac{1}{2}\left(\frac{\bar{q}}{2g}\right)^{1/2} \frac{t_c}{d'(j^*)} - \frac{1}{2}\gamma\left(\frac{\bar{q}}{g}\right)^{1/2} t_c}{\gamma g^{1/2}\bar{q}^{1/2} - rhs'\left(\frac{\bar{q}g}{2}\right)^{1/2} \frac{1}{d'(j^*)}} = \frac{\frac{1}{2}\left(\frac{\bar{q}}{g}\right)^{1/2} t_c}{(\bar{q}g)^{1/2}} \frac{rhs'\left(\frac{1}{2}\right)^{1/2} \frac{1}{d'(j^*)} - \gamma}{\gamma - rhs'\left(\frac{1}{2}\right)^{1/2} \frac{1}{d'(j^*)}} =$$

$$= \frac{1}{2}\frac{t_c}{g} \frac{rhs'\left(\frac{1}{2}\right)^{1/2} \frac{1}{d'(j^*)} - \gamma}{\gamma - rhs'\left(\frac{1}{2}\right)^{1/2} \frac{1}{d'(j^*)}} = -\frac{1}{2}\frac{t_c}{g} < 0$$

■

Proposition 3. *If storage costs increase, the average frequency of trips increases.*
Proof. Denote with

$$\bar{f} = (1 - T)\, \bar{f}^{(I=0)} + T \bar{f}^{(I=1)}$$

the average frequency of transactions in the economy. All arguments are implicit, T is the fraction of consumers traveling, $\bar{f}^{(I=0)} = \left(\frac{g\bar{q}}{2t_h}\right)^{1/2}$ is the average frequency of trips for those who do not travel, and

$$\bar{f}^{(I=1)} = \int_{t_h}^{t_h + t_c} \frac{1}{2}\left(\frac{g\bar{q}}{2t_h}\right)^{1/2} + \frac{1}{2}\left(\frac{g\bar{q}}{2t_s}\right)^{1/2} dT(t_s) < \bar{f}^{(I=0)}$$

is the average frequency of trips for those who travel. Denote with primes the same variables when g

increases to g'. Then

$$\bar{f} = (1-T)\,\bar{f}^{(I=0)} + T\bar{f}^{(I=1)} < \left(1-T'\right)\bar{f}^{(I=0)} + T'\bar{f}^{(I=1)}$$

since as g increases, less people travel and $T' < T$, and $\bar{f}^{(I=0)} > \bar{f}^{(I=1)}$. Moreover,

$$\left(1-T'\right)\bar{f}^{(I=0)} + T'\bar{f}^{(I=1)} < \left(1-T'\right)\bar{f}'^{(I=0)} + T'\bar{f}'^{(I=1)}$$

since $\bar{f}^{(I=i)} < \bar{f}'^{(I=i)}$ for $i \in \{0,1\}$. But $(1-T')\,\bar{f}'^{(I=0)} + T'\bar{f}'^{(I=1)} = \bar{f}'$, and hence,

$$\bar{f} < \bar{f}'$$

■

B Data Processing

B.1 Merchant codes

The transaction data classifies merchants using the MCC classification. Classification of merchants come at a "broad" and "narrow" level. We exclude narrow merchant categories that either refer to a transaction which can be executed without involving physical movement of a provider or a customer, or those that are of commercial, rather than private, nature. These categories broadly include items like airlines, cruise lines, direct marketers, online marketers, insurance, financial institutions, business services, political organizations, and other codes reserved for cash advances and balance transfers. The result is a classification of 27 broad categories, 221 nested narrow ones.

B.2 Transaction data

The raw transaction data comes from U.S. credit card statements issued between March and October 2003. Some earlier transactions still appear in the file as the date in which they are recorded may not necessarily be the date of the transaction. There are originally 3,530,027 records in the data for 134,008 unique accounts. Each record comes from a line in an individual credit card statement. A record contains the account number, transaction date and post date, amount and type of the transaction, the original merchant category code (MCC), and string information on the merchant name and location. After merging this data with the merchant codes above, 1,247,438 transactions are dropped. Of these dropped observation, around 1.1 million records are related to 1) cash advances, interests, late fees, account adjustments, balance transfers, card payments and similar activities not generated by actual purchases, 2) direct marketers and telemarketers, 3) unknown merchants. We further keep only records that are actual purchases ("transaction type" code equal to 253) originating on or after February 1, 2003. This leaves us with 2,207,907 transactions from 80,087 accounts.

B.3 Account data

The account data for the months of March to October 2003 has originally 2,272,825 records for 249,032 accounts. Among other things each line contains the record date (year and month) for the entry, the account number, a person ID, the date of birth and gender of the account holder, an external status code, a reported income, a 5 digit zip code and the State of residence. Different lines for the same account may be present in the account data because of various events that affect the account (the end of the billing cycle, updates to the month end balance, an income or residence change, for example). 28,928 observations appear to be of inactive cards (no information for State, zip code, and date of birth), so we drop them. Towards matching the account information with the transaction data, we start by keeping unique combinations of account, date of birth, state, zip code and record date. We find 4 accounts for which the date of birth of the account holder changes, and we make that information consistent by picking the oldest date of birth. After this adjustment, almost all records are unique within account number-event date. We drop three accounts, where the same set of several zip codes are reported for each record date making it difficult to find a residence location. This processing leaves us with 1,746,667 observations for 239,369 unique accounts.

B.4 Matching transactions and account data

We match the transaction and account data to assign a location of residence to each purchase. For a given account, we match the month of the transaction in the first file to the event month in the account data, if possible. For those observations where this is not possible, we match the closest account information that precedes the transaction; when this second option is not feasible, we match it with the earliest information following the transaction. The matching process leaves us with 2,189,048 transactions matched (more than 99% of the transactions data file) from 79,209 unique accounts. Out of the totality of matched transactions, only 155,254 did not find the exact event month in the account information: 145,815 records among these come from transactions in February 2003 (which are then matched with information in March).

B.5 Extracting merchant location name

The data provides us with a full merchant name string (including usually merchant name, location/phone number and State) and a merchant name string. Here we explain how we extract the potential city and State names of each transaction.

We first extract the merchant State. The State of the merchant is typically located at the end of the full merchant name. We extract the last two characters of the merchant name string if the last three start with a space. Only 6,147 transactions do not meet this requirement: inspection shows that in most cases, the last two letters still represent a State (or a foreign country), but we won't be able to rule out false positives. We match these States with a list of U.S. States and country abbreviations to verify that we have extracted U.S. States. We match only 86% of the 6 thousand problematic observations, and more than 98% of the other transactions. Keeping only transactions where a U.S. State could be identified leaves us with 2,154,927 observations.

To identify the set of observations that it may be possible to match with a location name, we start by extracting a potential location name. To do so, we remove from the full merchant name string the merchant name that the data provides (from the left of the string) and the State we have extracted (from the right of the string). This procedure generates 9,004 observations with an empty potential location name.

We then mark transactions of common online providers[20] and find the words "Online", "On Line", ".com", ".net" in 102,715 observations. We mark observations where the final part of the string before the State is a phone number (these are typically online stores) and find 206,012 of them. After this processing, we are left with 1,931,815 transactions that may contain city names, 90% of those for which a state name could be found, for 73,959 unique accounts. Note that the largest contributor to the drop

[20] We identify Paypal, QVC, AOL, Shutterfly, MUI Movies Unlimited, Amazon, Microsoft, Expedia, Untd.com, Ebay, Netflix.

in observations is transactions with a phone number (rather than a location) at the end of the merchant name. We will attempt to match this list of location names with a list of U.S. city and place names from the U.S. Census. Before turning to the different steps in that match, we will discuss briefly how we recover the list of cities.

B.6 List of cities and places in the United States

We construct a list of city names and States from the year 2000 U.S. Census Gazetteer list of Places and the year 2000 U.S. Census list of County Subdivisions. The list of places contains incorporated places and unincorporated Census Designated Places (CDP); it excludes towns in the New England states, New York, and Wisconsin, and boroughs in New York (treated as Minor Civil Divisions, or MCDs). The list of County Subdivisions contain, among other things, MCDs (called for example townships, parishes, districts), and Census County Divisions. Both lists contain, among other things, population in 2000 and latitude and longitude of the location.

While FIPS codes are unique, our match to merchants will be on a location name. Hence it may happen that within the list, we have more than one record with the same name (for example, we may have "Mountain View city" and "Mountain View, CDP"). In those cases, we attribute to a name the coordinates with the highest population in 2000.[21]

B.7 Finding location names in the transactions data and computing distances

We attempt to find the name of a city in four passes. First, we match the location name and State identified above with the list of U.S. Places. We immediately find a match for 1,475,545 out of the 1,931,815 we intend to match, 76% of our observations. Out of the 456,270 transaction with no match, 123,861 have names and states that match the MCD list. We assign "match quality" equal to 10 to those transactions matched at this first pass. We have 332,409 transactions with no location information (about 17% of the transactions) that we cannot match exactly.

In several instances, the name of a city in the Transaction data is truncated from the original. The second pass of the match involves matching truncated versions of city names from the U.S. Census to location names in the transaction data. We assign "match quality" equal to 9 to those cases where the name of a location in the transaction data, of length n, matches the first n characters of a city name. We further assign "match quality" equal to 8 where, for a location name of length n, there is a match in the first $n-1$ characters. Obviously, it can happen that one city in the transaction data can be matched to more than one city in the Census list. We only keep cases where the match is either unique or there are two matches. We solve the two-matches case as follows: if the match is to a Census place and to a minor civil division, we keep the coordinates of the Census place; otherwise, we take the place with the highest population and downgrade the "match quality" by 1. With the second pass, we are able to

[21] An alternative choice could have been to compute the average longitude and latitude of all the occurrences, weighted by population. However, we would still need a unique FIPS code identifier, since accounts will be associated to place codes, not names. This difference makes the approach infeasible.

recover 117,787 observations, bringing the number of matched transactions to 1,604,485.

In other instances, some locations may not be matched because of extra spaces, or special characters (e.g., "St. Louis" vs. "St Louis"). In the third pass, we "standardize" the name of the remaining unmatched locations by removing all spaces, commas, full stops and dashes both in the transaction and in the Census files. We assign "match quality" equal to 9 to these observations. With this process, we recover additional 21,080 observations, bringing the number of matched transactions to 1,738,273.

Finally, we identify the remaining unmatched locations with at least one thousand transactions and fix those matches by hand. There are 44 of these instances. We recover 33,700 observations more (also assigned "match quality" equal to 10), bringing the total to 1,771,973 matched transactions, or 91.7% of the transactions we intended to match. For these matched transactions we can attribute a latitude and longitude of the merchant.

The account data provides zip code information for each account. We match these zip codes against Census Places and (if we don't find a match) MCD lists using concordances for the year 2000 provided by the census. For the (few) cases in which we cannot find a correspondence, we use analogous zip-places and zip-MCD concordances for the year 2010. In some cases, a zip code may span two or more geographical units: we keep in that case the unit that accounts for the highest fraction of population of the zip code. We then have analogous geographies for account and merchant sides, and can compute the bilateral distance between the centroid of the account and shopping locations for each transaction.

The process of matching zip codes to geographical areas leads to a small loss in observations. Our working sample has 1,751,067 observations (90.6% of the transactions we intended to match) and 71,927 accounts. In our classification, 92.2% of observations have match quality equal to 10, and 7.2% have match quality 9, leaving less than 1% of observations with quality 8 (0.61%) and 7 (0.01%).

C Robustness

C.1 Summary Statistics by State

Table C.1 shows summary statistics on our main dataset by State of transaction.

C.2 Frequent users

Our data is unfortunately sparse enough not to allow a full analysis of individual consumers' behavior, since the median account uses the credit card around once per month, and around 96% of the accounts use the credit card less than once every two days. Here we focus on consumers with at least 15 transactions per month on average, and on transactions within 120 km from a consumer's residence. We term this as "frequent users" (FUs) sample, and use it to show that the limited mobility of consumers described above does not depend on including low frequency usage. Our FUs sample contains 2,198 accounts, conducting 496,654 transactions over the sample period. They reside in 1,729 locations and shop in 6,930 of them; there are a total of 26,436 origin-destination combinations over which we observe transactions.

Table C.2 shows summary statistics for such sample. Consumers in the median residence visit only 13 distinct sales locations overall during the sample period (15.3 sales location on average). Both values are higher than in the complete data; however, these consumers also live in places with richer options: the median residence here has 231 sales locations within 120 km (compared to 186 for the whole data). Hence, the median residence sees consumers shop in 5% of the available locations (the mean is 7%), very comparable to the values in the general data (4% and 6% respectively).

C.3 Percentiles of distances traveled

These Tables show summary statistics on the percentiles of distances traveled by consumers, by sector. Table C.3 refers to percentiles in the unweighted distribution. Table C.4 shows the same percentiles weighting each transaction with the correspondent purchase value.

C.4 Gravity over all distances

In Figure C.1, we estimate eq. (19) including origin-destination pairs at progressively longer distances. Specifically, we split all the (h, s) pairs in 20 quantiles of distances, and estimate it using only the first group, then only the first two, and so on, up to the whole set of observations. The blue line in Figure C.1 shows the coefficient on log distance. As one can see, changes of around +/- 30% in the 120 km cutoff (from 80 km to 160 km) only imply a variation in the gravity coefficient of around 0.1. Different sectors are more or less represented at different distances (see also Tables C.3 and C.4), implying that the coefficient δ varies.

Table C.1: **Summary of transaction amounts (in USD), by U.S. State of purchase**

State	Median	Mean	St. Dev.	Sum	N
AK	33	74	150	137,224	1,866
AL	28	64	173	1,090,956	17,060
AR	29	64	159	553,669	8,711
AZ	29	72	234	1,883,945	26,157
CA	31	74	213	11,058,192	149,493
CO	26	62	181	1,764,997	28,603
CT	31	71	191	4,238,450	60,117
DC	28	71	165	332,700	4,703
DE	30	73	217	494,720	6,743
FL	30	73	216	7,588,112	104,520
GA	27	64	184	2,720,214	42,260
HI	35	85	221	463,444	5,426
IA	28	61	170	814,099	13,454
ID	29	67	165	317,258	4,747
IL	30	69	184	4,797,273	69,414
IN	30	65	166	2,238,147	34,630
KS	29	63	186	1,000,398	15,786
KY	29	64	196	1,120,575	17,379
LA	30	63	146	1,201,510	19,111
MA	31	69	177	10,695,244	154,891
MD	28	69	190	2,491,372	36,154
ME	32	72	187	1,215,042	16,787
MI	30	67	173	3,257,719	48,499
MN	30	68	176	1,768,685	25,937
MO	29	67	191	1,937,271	28,953
MS	31	66	188	514,513	7,751
MT	33	69	169	305,490	4,424
NC	29	66	190	2,507,978	37,811
ND	29	65	149	220,194	3,387
NE	30	69	192	536,243	7,798
NH	32	81	277	1,873,150	23,081
NJ	32	73	213	7,459,863	102,554
NM	29	65	194	496,633	7,691
NV	41	93	238	1,241,557	13,377
NY	33	76	201	11,675,730	152,723
OH	29	66	176	3,843,913	57,932
OK	29	67	179	877,034	13,084
OR	28	64	187	1,278,495	19,997
PA	31	69	188	4,928,156	71,159
RI	31	70	176	1,157,634	16,559
SC	28	68	217	1,232,020	18,117
SD	34	74	217	247,896	3,349
TN	30	66	167	1,754,011	26,551
TX	26	62	184	6,154,744	98,592
UT	26	67	240	615,343	9,173
VA	28	68	201	3,038,906	44,872
VT	31	72	182	325,751	4,526
WA	27	66	195	1,548,626	23,514
WI	30	68	210	2,268,723	33,298
WV	32	68	177	393,698	5,779
WY	30	68	185	175,653	2,567
Total	30	70	195	121,853,167	1,751,067

Table C.2: **Summary statistics across residence locations - Frequent Users**

variable	min	p10	p25	p50	p75	p90	max	mean	N
Sales locations visited	1	5	9	13	20	27	139	15.29	1,729
Sales locations within 120km	7	80	137	231	494	838	1,111	344.2	1,729
Share available loc. visited	0	0.02	0.03	0.05	0.08	0.13	0.43	0.07	1,729

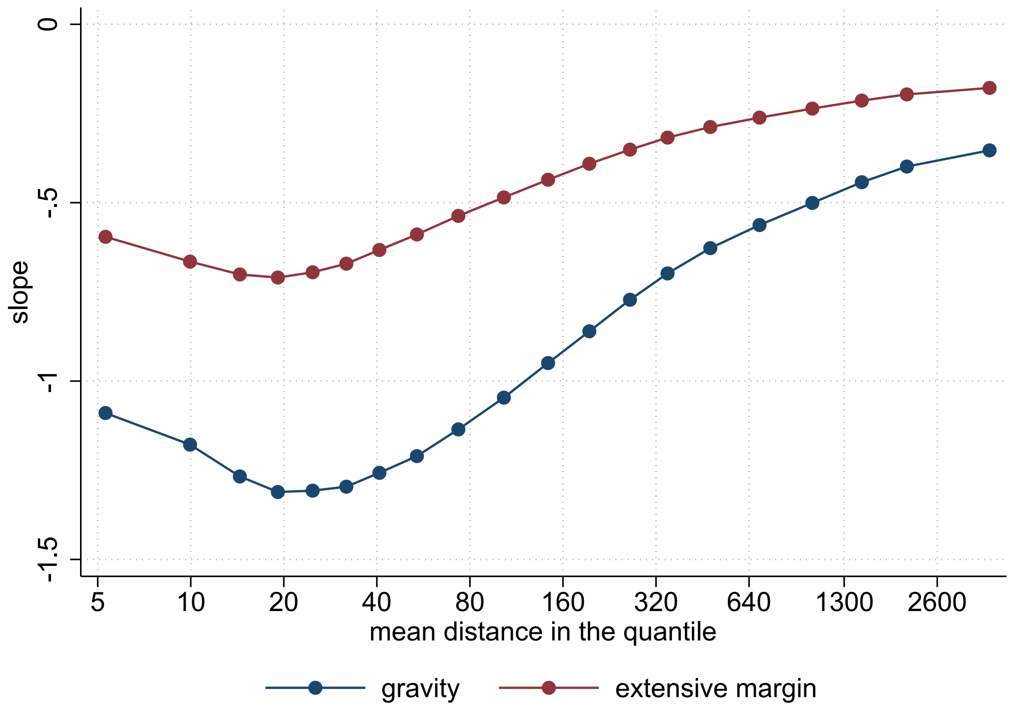

Figure C.1: Gravity in Expenditure

C.5 Margins decomposition

Figure C.2 shows the margins decomposition in eq. (21) applied to estimates of the gravity regression (19). The accounts margin is associated to 57% of the decline in expenditure in the typical sector as distance increases.

Tables C.5 shows the actual values of the account and expenditure margin with associated p-values represented in Figure 3; Table C.6 shows the actual values of the account and expenditure margin with associated p-values represented in Figure C.2

Tables C.7 and C.8 show the composition of frequency and batch size margin into the overall expenditure margin. They also show the share of the frequency margin in the expenditure margin, and the overall role of frequency and account margins in the total decline of expenditure with distance.

Table C.3: **Distribution of transaction distances (in km), by sector**

	p10	p25	p50	p75	p90	p99	max	mean
Agricultural Services	0	0	5.6	14.6	28.7	1,514.1	6,372.1	50.1
Amusement, Rec. Serv.	0	7.9	33.3	327.7	1,600.3	4,130	8,237.4	454.7
Apparel	0	4.7	15.6	52.1	364.7	3,833.6	8,253.1	201.4
Auto Repair/Service/Parking	0	0	7.9	24	78	2,315.3	7,937.3	94.8
Auto and Truck Sales/Service/Parts	0	0	8.3	21	58.6	2,119	7,775.3	88.8
Building Mat./Hardware/Garden Supp.	0	0	7.6	18.2	40.6	1,493.2	7,868.1	49.6
Communications	0	6.5	24.3	685.1	2,018	3,944.5	8,134.9	551.5
Durable Goods	0	5.7	21.8	161.8	1,650.7	3,946.5	7,115	419.5
Eating and Drinking Places	0	0	12.6	50.4	496.4	3,739.9	8,254.8	217.2
Educational Services	0	4.5	21.8	141.1	957.9	4,024.4	7,981.4	303.3
Food Stores	0	0	4.2	15.3	54	2,416.2	8,218	93.8
Furniture, Home Furnishings, Equip.	0	1.8	11.2	26.3	133.6	3,299.6	8,243.6	136.8
Gasoline Services	0	0	8.9	34.6	275.1	2,278.2	8,233.3	126.9
General Merchandise Stores	0	0	8.7	20.8	61.4	2,001.5	8,223.9	87.4
Health Services	0	0	8.6	20.3	46.3	2,231.1	7,969.9	83.9
Hospitality	51.3	162.8	367	1,011.8	2,257.9	4,158.5	8,253.1	801.9
Misc. Retail	0	0	8.6	29.6	355.1	3,736.8	8,223.9	193.1
Misc. Services	0	2.2	15.8	67.8	1,131.8	3,905.3	7,765.3	302.2
Motion Pictures	0	0	5.7	16.7	64.4	3,756.9	7,884.2	126.5
NonDurable Goods	0	0	8.2	22	143.9	3,437.6	7,768.4	145.4
Other Vehicles Sales/Service/Parts	0	6.6	20.1	55.1	506.1	3,018.8	7,879.4	191.5
Personal Services	0	0	6.7	20	135.1	3,332.5	8,251.4	132.3
Transportation Services	0	14.1	34.7	212.8	1,502.2	4,190.6	8,158.5	408.1
Vehicle Rental	0	6.8	79.9	1,506	2,813.7	4,558.8	8,216.8	893.7
Total	0	0	9.2	30.3	293.2	3,355.2	8,254.8	163.6

Table C.4: **Value-Weighted Distribution of transaction distances (in km), by sector**

	p10	**p25**	**p50**	**p75**	**p90**	**p99**	**max**	**mean**
Agricultural Services	0	0	6.7	16.8	36.3	1,348.7	6,372.1	52
Amusement, Rec. Serv.	0	8.3	37.5	418.8	1,752.7	4,290.3	8,237.4	530.2
Apparel	0	5.3	16.8	56.1	437.4	3,864.6	8,253.1	222.1
Auto Repair/Service/Parking	0	0	7.5	20.3	65.2	2,080.8	7,937.3	86.9
Auto and Truck Sales/Service/Parts	0	0	11.7	27.8	113.3	2,246.5	7,775.3	105.8
Building Mat./Hardware/Garden Supp.	0	0	9.9	23.4	54.2	1,577.2	7,868.1	56.3
Communications	0	4.5	14.7	113.6	1,518.2	3,818.2	8,134.9	367.8
Durable Goods	0	10.5	30.5	197.9	1,866	4,017.6	7,115	454.3
Eating and Drinking Places	0	0.8	15.4	79.2	708.8	3,940.4	8,254.8	264.2
Educational Services	0	8.6	24.6	107.8	632	3,913.2	7,981.4	262.4
Food Stores	0	0	5.2	16.9	55	2,378.1	8,218	91.9
Furniture, Home Furnishings, Equip.	0	4.6	13.1	30.7	129.7	2,970.8	8,243.6	129.9
Gasoline Services	0	0	9.7	39.6	320.1	2,252.9	8,233.3	133.9
General Merchandise Stores	0	0	9.9	23.1	77.2	2,555.2	8,223.9	104.1
Health Services	0	0	9.8	24.9	75.7	2,658.9	7,969.9	112.2
Hospitality	59.5	179.2	434.7	1,320.4	2,665	4,331.4	8,253.1	949.3
Misc. Retail	0	0	13	49.7	704.6	3,911.7	8,223.9	254.8
Misc. Services	0	5.3	17.2	54.7	666.7	3,964.9	7,765.3	238.8
Motion Pictures	0	0	7.3	22.2	223.1	3,960.8	7,884.2	181.9
NonDurable Goods	0	3	11.2	34.1	742.2	3,942.2	7,768.4	249.7
Other Vehicles Sales/Service/Parts	0	9.1	23.1	64.9	939	3,139	7,879.4	244.9
Personal Services	0	0	11	38.5	526.8	3,856.4	8,251.4	218.3
Transportation Services	4	16.9	77	651.5	2,004.8	7,346.6	8,158.5	683.3
Vehicle Rental	0	10.5	509.5	1,789.3	3,280.8	4,813.8	8,216.8	1,099.8
Total	0	0	12.8	44	521.3	3,822.1	8,254.8	224

Table C.5: **Expenditure out of home place (distances up to 120km)**

Category	Overall coeff	Overall pv	Accounts Margin coeff	Accounts Margin pv	Expenditure Margin coeff	Expenditure Margin pv	Share Accounts Margin	Obs.
Food Stores	-2.23	0.00	-1.12	0.00	-1.11	0.00	0.50	22,652
Gasoline Services	-2.08	0.00	-0.97	0.00	-1.11	0.00	0.47	39,673
General Merchandise Stores	-1.79	0.00	-1.08	0.00	-0.71	0.00	0.60	26,845
Misc. Retail	-1.70	0.00	-1.07	0.00	-0.63	0.00	0.63	34,057
Eating and Drinking Places	-1.57	0.00	-0.93	0.00	-0.64	0.00	0.59	34,509
Building Mat./Hardware/Garden Supp.	-1.40	0.00	-0.87	0.00	-0.53	0.00	0.62	14,190
Auto Repair/Service/Parking	-1.25	0.00	-0.88	0.00	-0.38	0.00	0.70	4,415
NonDurable Goods	-1.16	0.00	-1.05	0.00	-0.11	0.45	0.91	978
Health Services	-1.12	0.00	-0.77	0.00	-0.36	0.00	0.68	5,136
Apparel	-1.10	0.00	-0.83	0.00	-0.27	0.00	0.75	15,921
Transportation Services	-1.09	0.00	-0.68	0.00	-0.41	0.07	0.62	743
Furniture, Home Furnishings, Equip.	-1.07	0.00	-0.85	0.00	-0.23	0.00	0.79	12,292
Auto and Truck Sales/Service/Parts	-1.04	0.00	-0.81	0.00	-0.23	0.00	0.78	7,302
Motion Pictures	-1.04	0.00	-0.85	0.00	-0.19	0.00	0.82	1,927
Amusement, Rec. Serv.	-1.02	0.00	-0.66	0.00	-0.36	0.00	0.64	2,959
Educational Services	-1.00	0.00	-0.86	0.00	-0.13	0.56	0.87	712
Personal Services	-0.96	0.00	-0.89	0.00	-0.07	0.12	0.93	5,204
Vehicle Rental	-0.95	0.00	-0.83	0.00	-0.11	0.50	0.88	546
Misc. Services	-0.92	0.06	-0.63	0.00	-0.29	0.52	0.69	222
Communications	-0.89	0.00	-0.61	0.00	-0.28	0.04	0.69	424
Agricultural Services	-0.88	0.00	-0.66	0.00	-0.21	0.11	0.76	552
Other Vehicles Sales/Service/Parts	-0.68	0.41	-0.71	0.00	0.03	0.97	1.04	257
Hospitality	-0.65	0.01	-0.49	0.00	-0.17	0.35	0.75	1,394
Durable Goods	-0.09	0.90	-0.27	0.04	0.18	0.76	3.15	79

Table C.6: **Gravity in expenditure (distances up to 120km)**

Category	Overall coeff	Overall pv	Accounts Margin coeff	Accounts Margin pv	Expenditure Margin coeff	Expenditure Margin pv	Share Accounts Margin	Obs.
Food Stores	-0.85	0.00	-0.36	0.00	-0.50	0.00	0.42	18,635
Gasoline Services	-0.60	0.00	-0.25	0.00	-0.35	0.00	0.41	34,621
General Merchandise Stores	-0.93	0.00	-0.50	0.00	-0.43	0.00	0.54	23,933
Misc. Retail	-0.65	0.00	-0.40	0.00	-0.25	0.00	0.61	30,046
Eating and Drinking Places	-0.56	0.00	-0.31	0.00	-0.25	0.00	0.55	31,028
Building Mat./Hardware/Garden Supp.	-0.73	0.00	-0.39	0.00	-0.34	0.00	0.53	11,610
Auto Repair/Service/Parking	-0.40	0.00	-0.23	0.00	-0.16	0.00	0.59	3,014
NonDurable Goods	-0.65	0.00	-0.40	0.00	-0.24	0.01	0.62	758
Health Services	-0.33	0.00	-0.25	0.00	-0.08	0.09	0.75	3,914
Apparel	-0.53	0.00	-0.36	0.00	-0.17	0.00	0.67	14,069
Transportation Services	-0.47	0.00	-0.16	0.00	-0.31	0.00	0.34	635
Furniture, Home Furnishings, Equip.	-0.57	0.00	-0.40	0.00	-0.17	0.00	0.70	10,740
Auto and Truck Sales/Service/Parts	-0.33	0.00	-0.26	0.00	-0.07	0.08	0.79	5,508
Motion Pictures	-0.34	0.00	-0.28	0.00	-0.07	0.24	0.81	1,253
Amusement, Rec. Serv.	-0.22	0.00	-0.10	0.00	-0.12	0.00	0.45	2,330
Educational Services	-0.15	0.38	-0.19	0.00	0.04	0.83	1.24	530
Personal Services	-0.31	0.00	-0.27	0.00	-0.04	0.26	0.86	3,761
Vehicle Rental	-0.08	0.59	-0.22	0.00	0.14	0.30	2.71	296
Misc. Services	0.97	0.01	-0.10	0.07	1.07	0.00	-0.10	120
Communications	-0.41	0.01	-0.25	0.00	-0.15	0.22	0.63	263
Agricultural Services	0.42	0.11	-0.12	0.21	0.54	0.03	-0.28	190
Other Vehicles Sales/Service/Parts	-0.59	0.08	-0.07	0.17	-0.51	0.10	0.13	128
Hospitality	-0.14	0.08	-0.08	0.00	-0.06	0.38	0.55	1,160
Durable Goods	1.11	0.67	0.00	1.00	1.11	0.67	0.00	15

Table C.7: **Expenditure out of home place: number of transactions and average expenditure**

Category	Expenditure margin		Batch size margin		Frequency margin		Share of Frequency margin	Share of Account+Frequency margins	Obs.
	coeff	pv	coeff	pv	coeff	pv			
Food Stores	-1.11	0.00	-0.18	0.00	-0.93	0.00	0.84	0.92	22,652
Gasoline Services	-1.11	0.00	-0.09	0.00	-1.02	0.00	0.92	0.96	39,673
General Merchandise Stores	-0.71	0.00	-0.06	0.00	-0.65	0.00	0.91	0.97	26,845
Misc. Retail	-0.63	0.00	0.05	0.00	-0.68	0.00	1.08	1.03	34,057
Eating and Drinking Places	-0.64	0.00	0.02	0.05	-0.66	0.00	1.04	1.02	34,509
Building Mat./Hardware/Garden Supp.	-0.53	0.00	-0.02	0.45	-0.51	0.00	0.96	0.99	14,190
Auto Repair/Service/Parking	-0.38	0.00	-0.21	0.00	-0.16	0.00	0.44	0.83	4,415
NonDurable Goods	-0.11	0.45	0.02	0.88	-0.13	0.04	1.17	1.02	978
Health Services	-0.36	0.00	-0.17	0.00	-0.19	0.00	0.52	0.85	5,136
Apparel	-0.27	0.00	-0.01	0.53	-0.26	0.00	0.95	0.99	15,921
Transportation Services	-0.41	0.07	-0.26	0.16	-0.15	0.19	0.37	0.76	743
Furniture, Home Furnishings, Equip.	-0.23	0.00	-0.01	0.88	-0.22	0.00	0.97	0.99	12,292
Auto and Truck Sales/Service/Parts	-0.23	0.00	-0.01	0.85	-0.22	0.00	0.96	0.99	7,302
Motion Pictures	-0.19	0.00	0.02	0.69	-0.21	0.00	1.11	1.02	1,927
Amusement, Rec. Serv.	-0.36	0.00	-0.19	0.01	-0.18	0.00	0.49	0.82	2,959
Educational Services	-0.13	0.56	-0.11	0.63	-0.03	0.69	0.22	0.89	712
Personal Services	-0.07	0.12	0.16	0.00	-0.23	0.00	3.31	1.17	5,204
Vehicle Rental	-0.11	0.50	-0.05	0.75	-0.06	0.29	0.56	0.95	546
Misc. Services	-0.29	0.52	-0.20	0.65	-0.09	0.36	0.32	0.79	222
Communications	-0.28	0.04	-0.17	0.25	-0.11	0.09	0.38	0.81	424
Agricultural Services	-0.21	0.11	0.01	0.94	-0.22	0.00	1.04	1.01	552
Other Vehicles Sales/Service/Parts	0.03	0.97	0.27	0.71	-0.24	0.28	-7.96	1.39	257
Hospitality	-0.17	0.35	-0.05	0.77	-0.12	0.09	0.72	0.93	1,394
Durable Goods	0.18	0.76	0.02	0.97	0.17	0.50	0.91	1.19	79

Table C.8: **Gravity in expenditure: number of transactions and average expenditure**

Category	Expenditure margin		Batch size margin		Frequency margin		Share of Frequency margin	Share of of Account+Frequency margins	Obs.
	coeff	pv	coeff	pv	coeff	pv			
Food Stores	-0.50	0.00	-0.13	0.00	-0.36	0.00	0.73	0.84	18,635
Gasoline Services	-0.35	0.00	-0.04	0.00	-0.31	0.00	0.89	0.93	34,621
General Merchandise Stores	-0.43	0.00	-0.09	0.00	-0.33	0.00	0.78	0.90	23,933
Misc. Retail	-0.25	0.00	-0.01	0.16	-0.24	0.00	0.95	0.98	30,046
Eating and Drinking Places	-0.25	0.00	-0.02	0.00	-0.23	0.00	0.90	0.96	31,028
Building Mat./Hardware/Garden Supp.	-0.34	0.00	-0.07	0.00	-0.27	0.00	0.80	0.91	11,610
Auto Repair/Service/Parking	-0.16	0.00	-0.09	0.06	-0.07	0.00	0.43	0.77	3,014
NonDurable Goods	-0.24	0.01	-0.09	0.23	-0.15	0.00	0.62	0.86	758
Health Services	-0.08	0.09	0.03	0.53	-0.11	0.00	1.33	1.08	3,914
Apparel	-0.17	0.00	-0.02	0.11	-0.15	0.00	0.90	0.97	14,069
Transportation Services	-0.31	0.00	-0.09	0.24	-0.22	0.00	0.70	0.80	635
Furniture, Home Furnishings, Equip.	-0.17	0.00	-0.04	0.06	-0.13	0.00	0.77	0.93	10,740
Auto and Truck Sales/Service/Parts	-0.07	0.08	0.02	0.53	-0.09	0.00	1.33	1.07	5,508
Motion Pictures	-0.07	0.24	0.02	0.64	-0.09	0.02	1.32	1.06	1,253
Amusement, Rec. Serv.	-0.12	0.00	-0.04	0.33	-0.08	0.00	0.67	0.82	2,330
Educational Services	0.04	0.83	-0.01	0.94	0.05	0.37	1.33	0.92	530
Personal Services	-0.04	0.26	0.08	0.02	-0.12	0.00	2.79	1.25	3,761
Vehicle Rental	0.14	0.30	0.19	0.16	-0.05	0.28	-0.35	3.31	296
Misc. Services	1.07	0.00	1.19	0.00	-0.12	0.14	-0.12	-0.23	120
Communications	-0.15	0.22	-0.24	0.05	0.09	0.15	-0.61	0.40	263
Agricultural Services	0.54	0.03	0.68	0.01	-0.15	0.35	-0.27	-0.63	190
Other Vehicles Sales/Service/Parts	-0.51	0.10	-0.51	0.10	-0.00	0.98	0.01	0.13	128
Hospitality	-0.06	0.38	-0.05	0.39	-0.01	0.75	0.16	0.62	1,160
Durable Goods	1.11	0.67	0.96	0.73	0.15	0.79	0.14	0.14	15

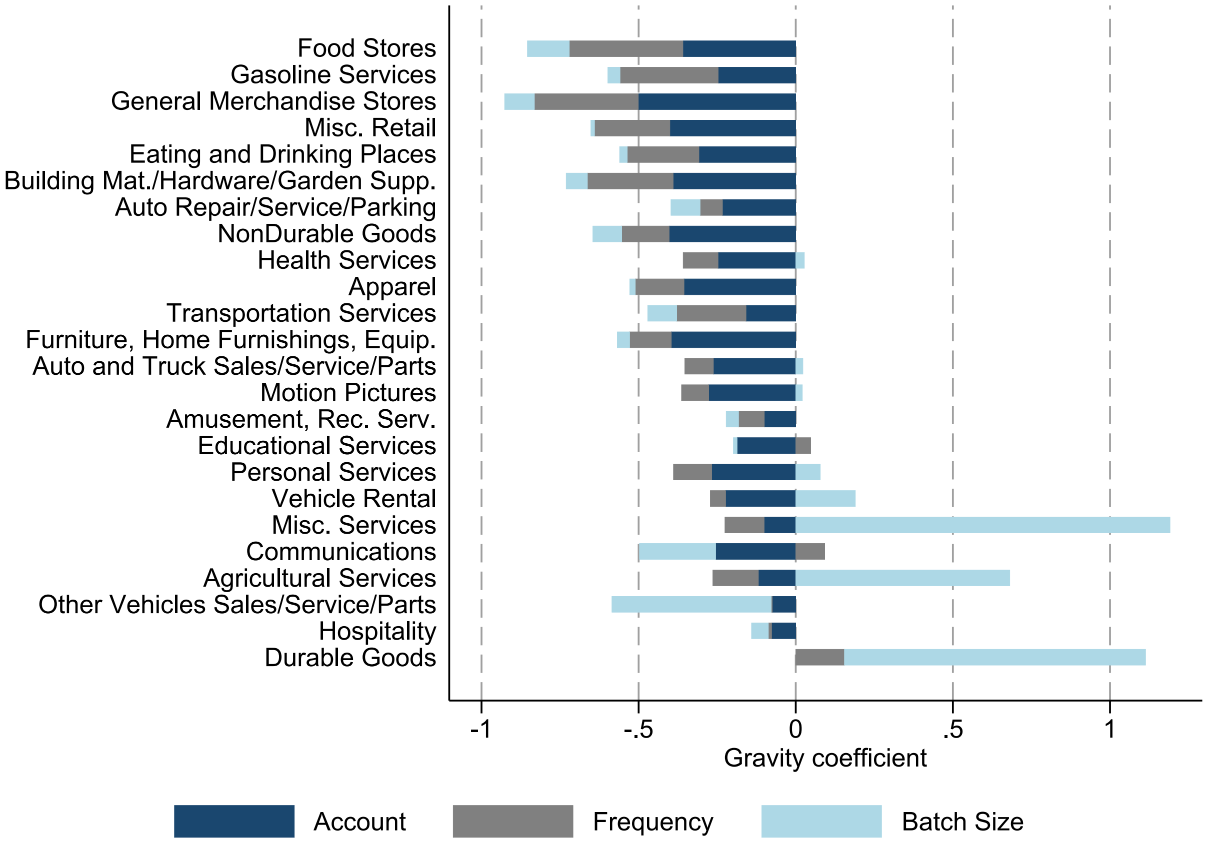

Figure C.2: Margins in gravity regression

C.6 Gravity and the frequency of transactions

These figures show further robustness on the relation between gravity and the frequency of transactions. Figure C.3 shows the correspondent of Figure 4 using all coefficients, and not just the ones significantly different from zero. Figure C.4 use the strength of gravity as measured by regression (19), only showing the coefficients significantly different from zero. Figure C.5 shows the correspondent of figure C.4, using all estimated slopes.

C.7 Rain and the flattening of gravity

Figure C.6 below parallels figure 5 in the main text, and includes all coefficients where the interaction between out-of-home expenditure and rain is significantly different from zero.

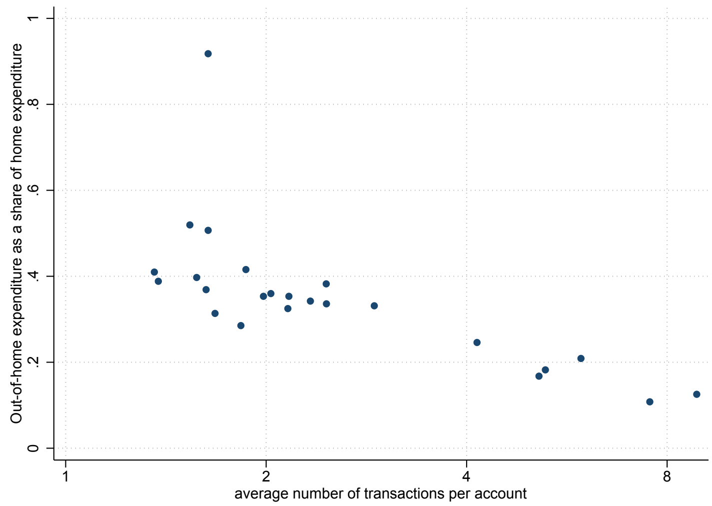

Figure C.3: Drop in expenditure out of home (all coefficients)

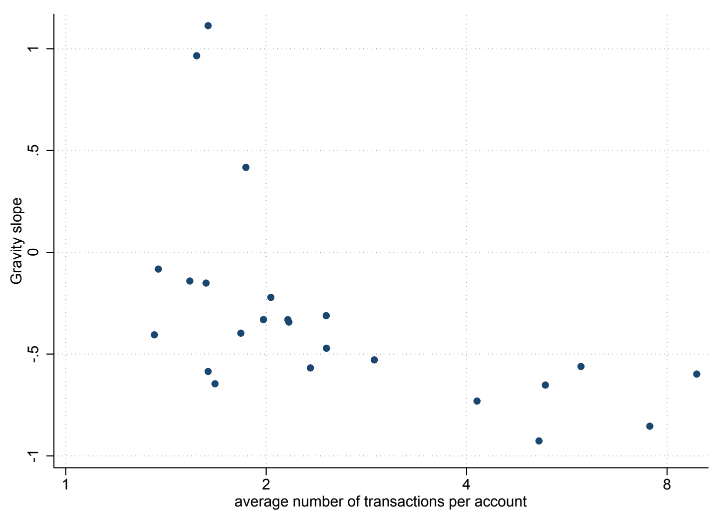

Figure C.4: Gravity and frequency of transactions (only significant slopes)

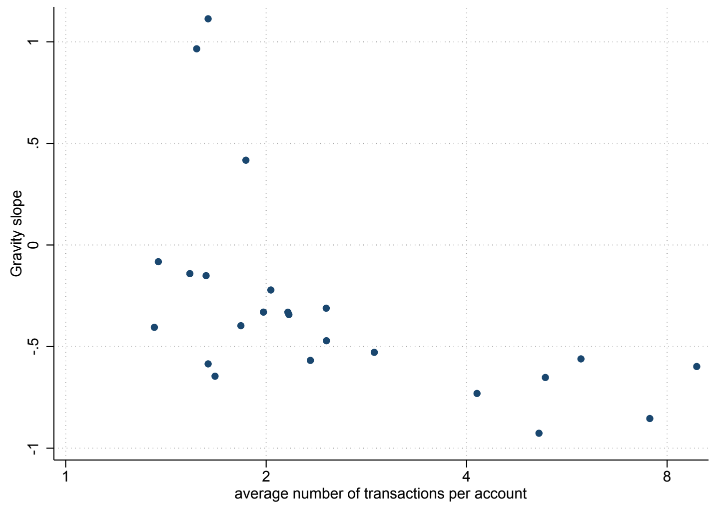

Figure C.5: Gravity and frequency of transactions (all slopes)

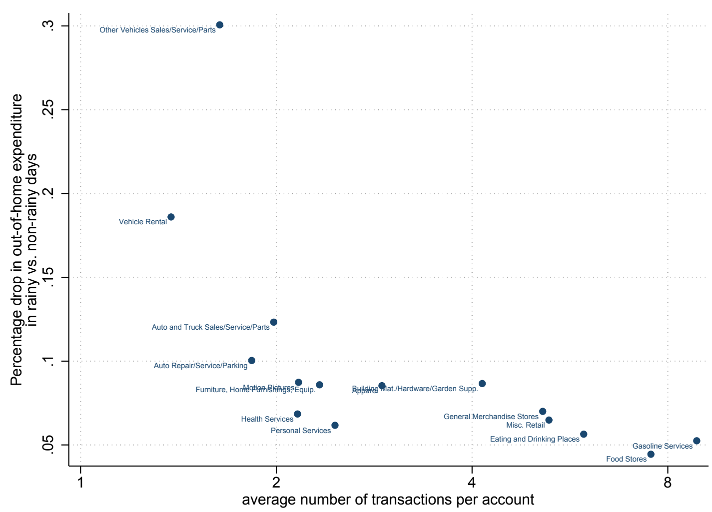

Figure C.6: The flattening in gravity across sectors