Shopping Distancing: The Impact of Travel Cost on Shopping Destination Choices

Albert Valentí
IESE Business School, University of Navarra, Av. Pearson 21, 08034, avalenti@iese.edu,

Victor Martínez-de-Albéniz
IESE Business School, University of Navarra, Av. Pearson 21, 08034, valbeniz@iese.edu,

Travel cost is a primary factor driving consumer choice of shopping destination. The literature has suggested that an increase in travel cost reduces shopping visits. However, we develop a choice model that predicts that with an increase in travel cost consumers tend to substitute shopping at distant venues for nearby options. That is, an increase in travel cost leads to an increase in shopping visits for some customers. We exploit the natural shock brought by the COVID-19 pandemic to empirically validate our theory, using population-wide data sets of mall visits from four cities in two countries. Our empirical results show that there is a threshold in the order of 500 meters from the shopping destination below which visits increased due to increased travel costs during the pandemic lockdown. During the reopening phase, the threshold becomes larger or smaller depending on the type of economic recovery. We provide further evidence that this phenomenon is driven by customer choices showing that ability to choose exacerbates the impact of travel cost. Our findings suggest that urban mobility restrictions, which increase travel costs, reduce visits from those living far away, but may lift demand from locals.

Key words: retail; travel cost; shopping choices; choice models; customer mobility; COVID-19

1. Introduction

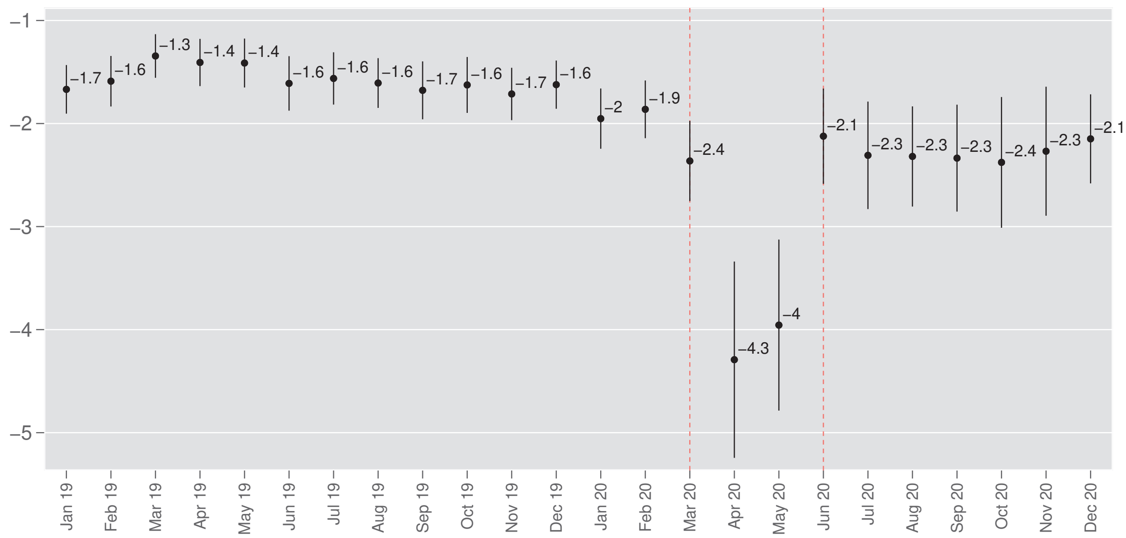
In the last decade, the future of physical retail has been put into question. From claims that America is 'overstored' (Kahn 2021) to the advent of online retailing (Caro et al. 2020), stores seem to be less relevant to consumers. Indeed, footfall has been falling in recent years (Nazir 2019), even before store closures during the lockdowns in the COVID-19 pandemic. At the heart of this trend is the realization that, while stores contribute with an experiential value into the shopping process, store visit also involves a travel cost that consumers may not always be willing to incur. When the travel cost is zero, visiting a store provides a higher utility compared to shopping online (Bell et al. 2020). In contrast, when travel cost is high, consumers prefer to avoid the visit and shop in lower access cost alternatives.

To help understand the future of retail and what shoppers will choose in future years, we quantify the sensitivity of consumers to travel cost and how this sensitivity may change. Unfortunately, travel cost is intrinsically connected to the geographical locations of customers and stores, which means that it is hard to empirically identify the impact of this cost on shopping propensity. Moreover, in measuring sensitivity to travel cost, it is important to properly identify the shopping options that consumers have. These options are not only about the offline vs. online channel choice (Chintagunta et al. 2012), but more broadly about the different physical locations that compete for consumers' attention. Only through this broader choice set, we can fully capture the choice process that consumers face and the role of travel cost.

To this end, we take advantage of a major natural experiment, the COVID-19 pandemic, to empirically measure the impact of travel cost on consumer choices across different shopping options. Indeed, COVID-19 has radically changed the way we shop, not only in the short-term during store closures in Spring of 2020, but also afterwards with mobility restrictions that have made store visits more costly. In this sense, we take the pandemic as a shock on travel cost, driven by a restricted mobility and a psychological burden from the risk of infection when travelling. We recognize that the pandemic has provoked a shock on both demand and supply factors. Hence, the main effect of the pandemic shock captures many of these factors (e.g., product availability). However, the travel cost for individuals within the same city captures only the demand factor effect of the distance to shopping destination.

To illustrate the strength of the COVID-19 shock on travel cost, we specify the simple linear regression in Equation (1) to evaluate how the relationship between visits and distance to shopping destination has evolved. The regression evaluates how the relationship between the number of visits

Figure 1 **Evolution of the Relationship between Distance to Mall and Visits (Mall in Barcelona)**



Notes. The red-dashed lines indicate the period of lockdown. Vertical axis represents the coefficient and confidence interval estimates of β_t in Equation (1).

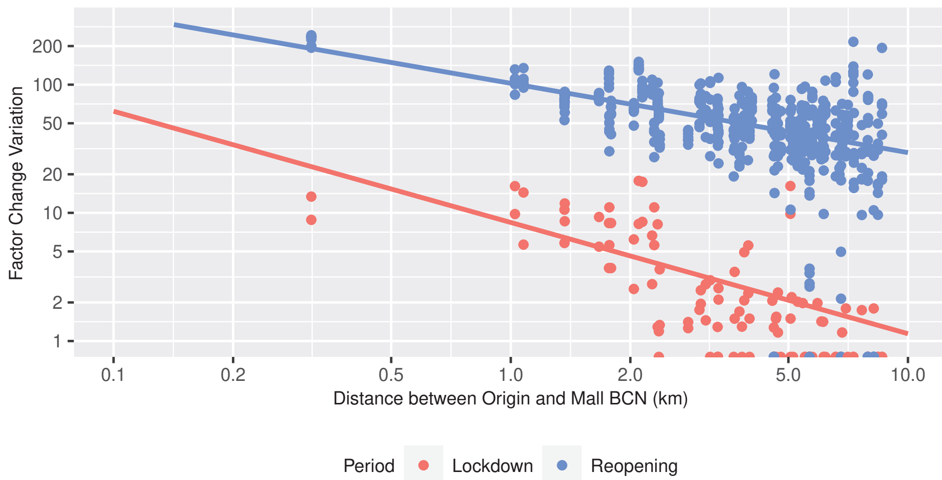
to a shopping mall from district i in month t and the distance from the district to the mall evolves over time, controlling for seasonality and district population. We estimate the regression with a population-wide data set that geolocates visits from all individuals in the city of Barcelona, Spain, from January 2019 to December 2020.

$$\log(visits_{it}) = \alpha_{0t} + \beta_t \log(Distance_i) + \gamma_t \log(Population_i) + \varepsilon_{it} \tag{1}$$

Figure 1 shows the estimates of the distance coefficients, β_t. During the most severe lockdown from March to May 2020, the coefficient β_t was sharply reduced, indicating a stronger travel cost than before the pandemic. After June 2020, the β_t coefficient recovered but stabilized to a lower level compared to pre-COVID, indicating an increased cost of mobility during this period of reopening compared to pre-COVID.

To provide a more nuanced model-free evidence of the COVID-19 shock on travel cost, we plot in Figure 2 the variation in monthly visits to the shopping mall before vs. after COVID-19 for each district. We emphasize multiple aspects. First, the sharper decline in visits during lockdown

Figure 2 **Model-free Relationship between Distance to Mall and Variation in Visits by District (Mall in Barcelona)**



Notes. Each dot represents the factor change variation in visits of a district-month. Factor change variation is computed by dividing the number of daily visits from each month from March to December of 2020 by the corresponding month in 2019 multiplied by 100. A number above 100 indicates an increase in the number of visits and a decrease otherwise. The two lines are the linear OLS predictions of the relationship between distance and variation in visits.

compared to the reopening phase noted in Figure 1 occurs across districts, since almost all red dots are below the blue dots. Second, not only there is an overall decline in visits post-COVID-19, but there is a negative relationship between distance to the mall and drop in visits which is stronger in the lockdown compared to the reopening period, as indicated by the steeper red trend vs. the blue trend. Taking both results together, the shock on travel cost is related to a main reduction in visits and an increased reduction of visits with distance. Third, and most notably, the trend line in the reopening phase intersects with the 100-factor-line indicating that for short distances to the mall the number of visits would increase with respect to pre-COVID-19 period. In fact, districts near the mall have experienced an increase in visits in the reopening period, as shown by the dots above the 100-factor-line.

Of course, Figures 1 and 2 do not constitute evidence of any general behavior nor that an increase in travel cost would cause an increase in the propensity to visit of customers close to the shopping destination. Yet, the figures suggest that existing literature which typically assumes or implies that increases in travel cost always reduce the propensity to shop (e.g., Bell et al. 1998, Marshall and Pires 2018), are not consistent with what we observe here. Hence, these figures provide a compelling reason to explore whether this behavior is generalizable and to develop a new theoretical model for consumer choices over shopping options that explain why nearby shopping destinations might benefit from higher travel cost.

For this purpose, we develop a model based on the Multinomial Logit (MNL) to characterize the probability of a given consumer choosing a shopping destination based on consumer and shop characteristics, time effects, and the distance between consumer origin and shopping destinations, which contribute linearly to travel cost. Namely, we include in our choice model insights from the existing literature on gravity models in trade economics, information economics, retail, and marketing. Our analytical model predicts that increases in the per-km travel cost can lead to increases in propensity to shop for destinations that are close to the individual.

We use the COVID-19 shock to identify the effect of changes in travel cost in our model. Specifically, we quantify how the increase in per-km travel cost associated with the pandemic has affected shopping destination choices depending on the distance from the district of origin. We employ geolocation data on shopping mall visits from four cities in two countries with very different shopping contexts: one in Latin America, where shopping malls are the main shopping option; and another in Europe, where street stores attract more customer attention compared to malls. We find that indeed the per-km travel cost has increased during the pandemic, leading to a concentration of shopping activity closer to each consumer's residence. Our model is able to explain consumer choices better than models that ignore competition and results are robust to alternative, simpler specifications (Ordinary Least Squares and logit).

The shock on travel cost brought by COVID-19 has had different intensities over time (differences between strict lockdown and reopening period), over space (we study two different countries, one in Latin America, the other in Europe), and consumer contexts (weekday vs. weekend shopping, weather conditions, and socioeconomic factors). We exploit the variations in consumer contexts to provide further evidence that the phenomenon we unveil is driven by the effect of travel cost on consumer choice. Namely, we show that in weekends, when there are fewer work obligations and consumers have more freedom to choose where to shop, travel cost increases more than during weekdays. Similarly, during good-weather days, when travel choice is less restricted than in rainy days, travel cost indeed increases more. Finally, affluent individuals who have higher freedom to choose when and where to shop experience higher travel cost compared to less-privileged individuals.

Our research thus advances our understanding of consumer choices in the earlier stages of the purchase funnel, informs policy makers and marketers on the effect of policies that affect mobility, and provides insights for the design of store networks. Indeed, our model and empirical results help predict how future changes in travel cost influence store footfall and customer mix by attracting more or less customers from different origins. The survival of urban stores is at the heart of the political debate on the future of dense urban areas (Hu 2016, Benvenuty 2021). Our results help policy makers understand the externalities of city planning decisions and mobility restriction policies. Specifically, our model predicts that visitors from suburban origins will sharply fall, because they substitute city-center destinations for their local store; in contrast, local residents will find it harder to "escape" the city and thus opt for local options. Anecdotically, when the city center of Madrid, Spain, was closed to motor vehicles during 2019-2020, newspapers shared evidence that customer profile of local stores had changed due to an increase in visits of nearby customers and a decrease of distant customers (Aranda 2019). Finally, our results also inform retail network design, providing a modeling tool to help evaluate the appropriate store density to attract consumers without excessive store costs.

2. Literature Review

2.1. Travel Cost and Shopping Choices

How consumers make shopping destination decisions is a topic of primary interest in marketing and retail. Marshall and Pires (2018) build a model of consumer grocery shopping that considers the decisions of where to shop and which products to shop. When making these decisions, consumers trade off between the travel cost and the prices and product assortment of each shop. The authors exploit the variation on weather, traffic, and cost of time to identify how travel cost affect store choices for a sample of 7,000 households. Their findings highlight the importance of travel cost on shopping decisions since travel cost rather than prices or assortment drives store choice. Bell et al. (1998) propose that consumers minimize total shopping cost and, hence, develop a store choice model that quantifies fixed and variable shopping costs. Fixed cost includes store preference and travel cost, which is modeled as a non-linear function of distance and varies by consumer. In their empirical application with 520 households, the authors find that while some segments of consumers are highly sensitive to travel cost, other segments are very insensitive. Chintagunta et al. (2012) develop a channel (online vs. offline) choice model that accounts multiple transaction costs. Travel cost consists of distance and time costs. To incorporate different effects of travel cost, distance is interacted with household income and distances is scaled by a factor of two in downtown areas and during peak hours. In their application with 3,500 customers of a retailer, higher travel cost encourages households to visit the online channel. Similarly, Forman et al. (2009) exploit the openings of physical stores to examine the effect of distance to store on online shopping. The authors find that both travel cost and the disutility of purchasing online influence consumer choices.

In sum, the literature predicts that an increase in travel cost would reduce shopping visits. Our work differs from the literature in that we analytically predict and empirically show that an increase in travel cost does not necessarily lead to a negative effect on shopping destination choices. Specifically, an increase in travel cost decreases the likelihood that a consumer chooses to visit a shopping destination that is far but increases the likelihood for a destination that is near. To provide further evidence that the phenomenon we unveil is driven by customer choices, we examine whether ability to choose moderates the travel cost effect. We do so, in line with previous studies, incorporating heterogeneity in the effect of travel cost in the three dimensions of weekday vs. weekend variation, weather conditions, and consumer social class.

2.2. Gravity Models

The law of gravitation has been widely used to model consumer choices of retail locations. Early conceptualization and formulation work goes back to Reilly (1931), Converse (1949), and Huff (1964). Multiple studies extended the gravity model proposing the inclusion of factors that influence consumer choice. Gautschi (1981) suggests the addition in the model of retail center characteristics such as assortment, design, and pricing as well as mode of transportation characteristics such as safety, comfort, and cost. Okoruwa et al. (1988) propose the inclusion of shopper demographic attributes and retail center attributes such as various size measures. Lee and Pace (2005) develop a gravity model that accounts for spatial dependencies among consumers and retailers. Li and Liu (2012) show that incorporating spatial competition and agglomeration improves the performance of the gravity model.

Building on these models, another stream of research provides approaches for retailers to select store locations by applying the law of gravitation to model the probability of consumers to choose among different retail locations (e.g., Ghosh and Craig 1983, Rust and Donthu 1995). Furthermore, other papers apply gravity models to examine online shopping choices to find that travel cost is also present online. These findings unveil that travel cost not only captures physical distance but also captures, especially for cross-regional shopping, differences in taste (Blum and Goldfarb 2006), culture (Burtch et al. 2014), regulation (Hortaçsu et al. 2009), infrastructure (Gomez-Herrera et al. 2014), reputation (Chintagunta and Chu 2021), and search costs (Lendle et al. 2016).

Our geolocation data and gravity model application have some parallelism with other population-level studies. The law of gravitation characterizes the traffic flow across the 30 main cities in South Korea (Jung et al. 2008) and the phone call patterns between cities for 2.5 million people in Belgium (Krings et al. 2009). In their seminal paper, Simini et al. (2012) propose a modelling framework based on gravitation, which they call radiation model, to predict population mobility patterns that only require information on population distribution. Beiró et al. (2018) use the gravity model to investigate whether social mixing affects the decision of 380,000 people to visit 16 malls in Chile.

Jia et al. (2020) document with the gravity model how the mobility of more than 11 million people in China predicts the spread of COVID-19 and the efficacy of mobility restrictions.

In sum, our paper builds on the above studies applying the gravity law to develop a new model of consumer shopping destination choices. Specifically, we embed the law of gravitation in the logit framework (McFadden 1974) so that shop attractiveness is proportional to distance to the power of $-\beta$. This model allows us to make analytical predictions on the effect of travel cost. We test empirically our predictions with customer data on daily visits during 27 months to seven malls in four cities from two countries for a total population of 5.4 million individuals.

2.3. COVID-19 Impact

Many recent studies examine how consumer spending habits changed due to the COVID-19 employing large data sets of credit card records. Carvalho et al. (2020) analyze 2.1 billion transactions from 6 million cardholders of a bank in Spain. These authors find a V-shaped consumption pattern which varies by province, sector, and channel (online vs. offline). They also find that mobility reduction during the lockdown is influenced by social class and day of the week, with poorer households traveling more during weekdays. Bounie et al. (2020) use 4.5 billion transactions from 70 million cards for all banks in France. They find that cardholders during lockdown reduced the distance traveled to one-quarter and concentrated spending in fewer retailers, a finding consistent with our model and empirical results. Relihan et al. (2020) analyze 450 million transactions from 11 million cardholders in US. They find that lower-income neighborhoods had a larger decline in overall spending and had a slower adoption of online grocery shopping although larger use of online restaurant ordering. Chronopoulos et al. (2020) utilize 20 million transactions from 100,000 cardholders in Great Britain. The authors identified stockpiling behavior and differences in consumption across sociodemographic groups.

Other studies employ different types of data to examine the heterogeneous impact of COVID-19 on consumption. For example, Chetty et al. (2020) combine data from private companies in the USA to track multiple economic indicators at the zip code, industry, income group, and business size level. They find that high-wage individuals faced a V-shaped recession that only lasted a few weeks, while low-wage individuals suffered much larger job losses that persisted several months. Similarly, Campos-Vazquez and Esquivel (2021) use point-of-sale transaction data from Mexico to find geographic and sectorial differences in consumption declines. Alexander and Karger (2020) combine credit card spending data with county-level data on mobility and mobility restrictions to find that political affiliation is associated to differences in mobility and restaurant spending. Related to changes in mall patronaging, He et al. (2020) examine the traffic evolution of 463 malls in China. The authors find that heterogeneity in foot traffic recovery after COVID-19 can be explained by pandemic situation and city characteristics such as population, GDP, and industrial structure. Martínez-de-Albéniz et al. (2021) break down the impact of COVID-19 on fashion retail sales into regulatory effects, social panic, and shocks from tourism. Alé-Chilet et al. (2020) document that the shock brought by COVID-19 induced a change in mobility patterns that explain a significant portion of the overall drop in non-respiratory emergency room visits during the period.

In line with these studies, we employ the COVID-19 outbreak as a major shock on consumer behavior. We argue that COVID-19 caused a shock on the travel cost faced by consumers due to both the mobility restrictions imposed by the authorities and the health risk associated with social interactions. In our empirical application, we account for differences in the strength of the shock, which we expect to peak during the lockdown and fade away afterwards. We incorporate individual and temporal heterogeneity examining differences by social class, weekday vs. weekend variation and weather conditions.

3. Theoretical Analysis

In this section, we provide a theoretical analysis of the effect of travel cost on customer shopping destination choices. We first present a general model and then derive our theoretical prediction from this stylized model.

We embed the travel cost customers face into a discrete choice model framework (McFadden 1974). This is similar to papers in economics of migration that have built on the logit model to generate a gravity model of destination choices (e.g., Beine et al. 2009, Grogger and Hanson 2011). Indeed, Anderson (2011) states that "the discrete choice probability model rationalizes structural gravity equally well" (p. 148). Hence, we express the utility obtained by individual i from choosing the shopping destination j as:

$$U_i^j = \alpha_i^j + \beta log\_dist_i^j + \epsilon_i^j$$

where the individual utility obtained from traveling to a certain shopping destination depends linearly on the individual preference for the destination (α_i^j), a (dis)utility from traveling to the destination (β, which should be negative to reflect that higher distance increases travel cost and reduces utility), and a residual term (ϵ_i^j). We express the distance variable in logs ($log\_dist_i^j$) for consistency with the gravity literature to account for a concave relationship.

When the residual follows an independent and identically distributed (i.i.d.) extreme-value distribution, we can apply the results in McFadden (1974) to write the destination choice probability as:

$$P_i^j := Pr(Y_i^j = j) = \frac{exp(\alpha_i^j + \beta log\_dist_i^j)}{\sum_{k=1}^{J} exp(\alpha_i^k + \beta log\_dist_i^k)} = \frac{a_i^j}{a_i^1 + a_i^2 + ... + a_i^J} \tag{2}$$

where $a_i^j = exp(\alpha_i^j + \beta log\_dist_i^j)$.

Next, we evaluate how a change in travel cost would influence destination choices. Taking logs in both sides of Equation (2), we obtain

$$\log(P_i^j) = \alpha_i^j + \beta log\_dist_i^j - \log\left(a_i^1 + a_i^2 + \ldots + a_i^J\right).$$

By taking derivatives with respect to β, we find that

$$\frac{\partial log(P_i^j)}{\partial \beta} = \frac{1}{P_i^j}\frac{\partial P_i^j}{\partial \beta} = log\_dist_i^j - \frac{(a_i^1 + a_i^2 + ... + a_i^J)'}{a_i^1 + a_i^2 + ... + a_i^J},$$

where $(a_i^j)' := \frac{\partial a_i^j}{\partial \beta} = a_i^j log\_dist_i^j$. Hence, we can write

$$\frac{\partial log(P_i^j)}{\partial \beta} = log\_dist_i^j - \sum_{k=1}^{J} w_i^k log\_dist_i^k$$

where $w_i^k = a_i^k / \sum_{k'=1}^{J} a_i^{k'}$ are positive weights that add to one. Hence, the probability of option j increases with respect to β if and only if its log-distance is higher than the weighted average of log-distances across other possible choices. In other words, as travel cost increases, $\beta < 0$ becomes even more negative, and then options that are nearby (lower than a threshold) see the number of visits increase, while those further away become less popular. In particular, $\frac{\partial (P_i^j)}{\partial \beta} < 0$ if destination j is the closest to individual i. This is formally stated next.

Theorem 1. *An increase in travel cost increases the propensity to visit the shopping destination that are closest to the individual and reduces the propensity to visit the furthest ones.*

4. Main Empirical Analysis: City of Quito

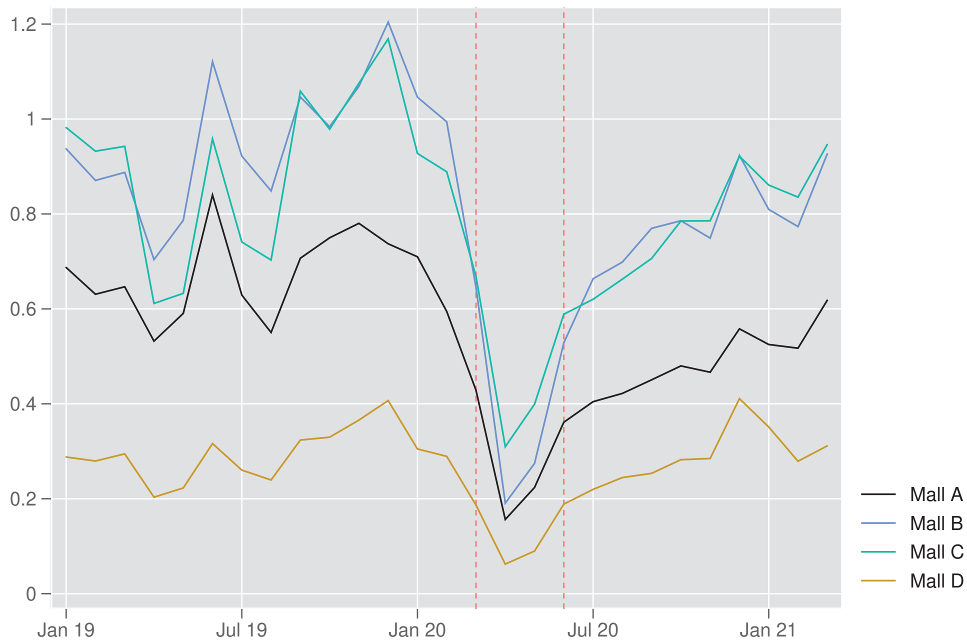
4.1. Data Description

Our primary data set was granted by a retail analytics company. The data geolocates individually via cellphone all the population in the city of Quito, Ecuador. For privacy reasons, individuals are aggregated into districts depending on where they live (in Ecuador called "subparroquias" (parroquias urbanas), "parroquias" (parroquias rurales), and "cantones"). The data tracks the number of visitors from each district to four main shopping malls in Quito from January 1, 2019, to March 31, 2021, for a total of 816 days.[1] We call the four malls A, B, C, and D. The number of unique visitors is measured at various time-aggregation levels ranging from daily to yearly. For our analysis, we consider the 51 districts with consumers visiting the four malls.[2] These districts account for a population of 2,544,382 individuals, with a minimum of 3,224 people in Tababela district and a maximum of 173,752 people in Calderon district. Online Appendix EC.1. locates the 51 districts in the city map of Quito and provides their population (variable *population*). Figure 3 shows the evolution of monthly visits per mall (variable *visit*).

[1] Five days from 2019 are missing from the data set.

[2] We excluded from the analysis 10 districts that due to the combination of small population size and large distance to the malls do not have any visitor to some malls throughout the whole span of the data.

Figure 3 **Evolution of Monthly Visits in Quito**



Notes. The red-dashed lines indicate the period of lockdown. The vertical axis reports the monthly visits in millions.

We distinguish three time periods in the data (categorical variable *Period*). The first period covers the pre-COVID-19 phase, the second period covers the lockdown, and the third period covers the reopening phase after the lockdown. In Quito, the lockdown started on March 17, 2020, and finished on June 3, 2020 (Gobierno de Ecuador 2020). During this period, the authorities imposed strict mobility restrictions across provinces and visits to malls were restricted to grocery shopping. From the end of the lockdown all types of shops could open with some capacity restrictions. We also call Periods 1, 2, and 3 as pre-COVID-19, lockdown, and reopening periods, respectively. Please see Online Appendix EC.2. for robustness checks with alternative period measurements.

The data set also contains the number of visitors of each district disaggregated by social class. To approximate the share of population that belongs to each social class group per district, we average the monthly share of visitors for each of the groups for the mall with highest number of visitors for the district. We focus on the wealthiest of the four social class group, which has a median and mean share of 18.0% and 17.4%, respectively. We code social class as a district-level time-invariant dummy variable that takes value of 1 if the share of the wealthiest social class of the district is above the median (dummy variable *rich*).

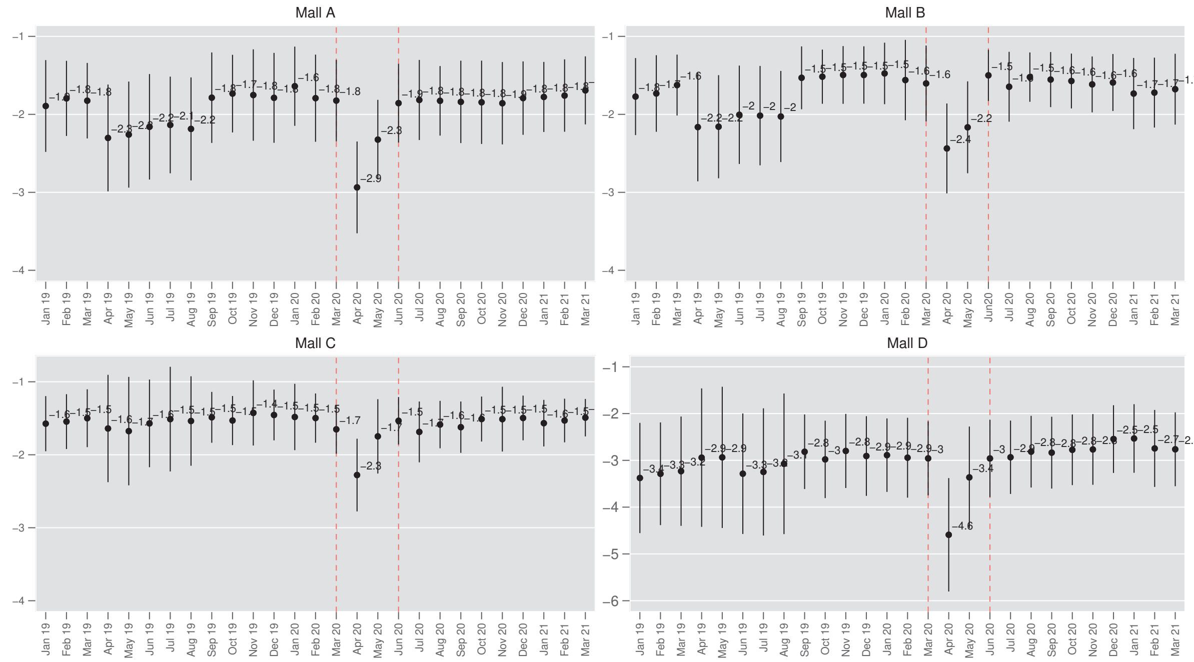
Finally, we collect data from multiple open-source platforms. We obtain the coordinates of the perimeter of each district from the city council of Quito, which we use to compute, for each district, the centroid and the distance to each mall; we take the log of distance to compute the variable *log_dist*.[3] The shortest and longest distance from a district (its centroid) to a mall are 1.21 and 37.73 km, respectively. We also collect the geographic location of the 10 malls in Quito not included in our primary data set. We compute which of the 14 malls is the closest to each district (categorical variable *closest*). We gather precipitation data from the NASA (NASA POWER Project 2003). The distribution of precipitation is highly skewed to the right. Although only 21 days have zero precipitation, 219 days have less than 1 millimeter (mm) of rainfall a day. The median and mean daily precipitations are 2.82 and 4.20 mm, respectively. We code precipitation as a dummy variable that takes value of 1 if the daily precipitation is below the median (dummy variable *no_rain*). Finally, we create a dummy variable that takes value of 1 on weekend days (variable *WE*).

4.2. Model-free Evidence

Before embarking on the model estimation, we show some model-free evidence of how changes in travel cost affect shopping destination choices differently depending on the distance, as predicted by our analytical model. We replicate for the four malls in Quito the two analyses presented in the introduction. First, we estimate the simple cross-sectional linear regression presented in Equation (1). In Figure 4, we plot the coefficient of *log_dist* for each month. Clearly, across malls, the cost of travel remains stable before the lockdown and suffers a shock during the lockdown, especially for Mall D. The cost of travel seems to return to pre-COVID-19 levels after the lockdown. Next, in Figure 5, we plot the relationship between distance to mall and change in number of monthly visits

[3] We consider alternative measurements for the distance variable. See Table 2 for non-parametric measurement. See Online Appendix EC.2. for a measurement in levels. Finally, distance measured in car driven time and car driven distance have correlations of .97 and .93 with our measure of bird's eye distance.

Figure 4 Evolution of the Relationship between Distance to Mall and Visits (Malls in Quito)



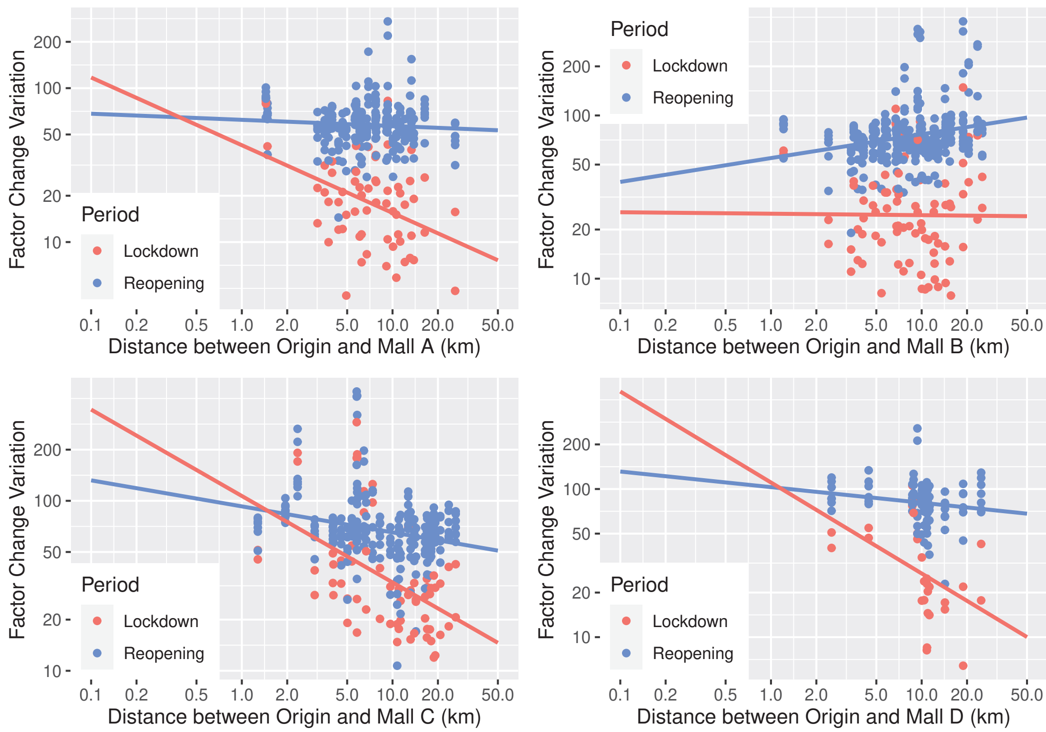
Notes. The red-dashed lines indicate the period of lockdown. Vertical axis represents the coefficient and confidence interval estimates of the variable β_t in Equation (1).

before and after COVID-19. Each dot represents the factor change in visits of a district-month. Similar to the opening analysis for the mall in Figure 2, we observe three main trends related to the shock on travel cost. First, the drop in number of visits increases with the distance to the mall, except in the post-lockdown period for Mall B. Second, this negative relationship between distance to the mall and drop in visits is stronger in the lockdown period when mobility restrictions were stricter. Third, the trends for Malls A, C, and D intersect with the 100-factor-line suggesting that for short distances the number of visits would increase post-COVID-19.

4.3. Model Formulation and Estimation

In this section we empirically test our prediction that an increase in travel cost increases the likelihood that a consumer chooses to visit a shopping destination that is near but decreases the likelihood for destinations that are far. We accommodate the choice model presented in Equation (2) to account for changes in travel cost at different periods in a panel framework. The following model presents the individual level decision to visit a shopping destination on a certain day:

Figure 5 **Model-free Relationship between Distance to Mall and Variation in Visits by District (Malls in Quito)**



Notes. Each dot represents the factor change variation in visits of a district-month. Factor change variation is the number of visits from each month from March 20 to December 2020 by the corresponding month multiplied by 100. A number above 100 indicates an increase in the number of visits and a decrease otherwise. The two lines are the linear OLS predictions of the relationship between distance and variation in visits.

$$P_{it}^{j} := Pr(Y_{it}^{j} = j) = \frac{exp(\alpha_d^j + \gamma_t^j + \delta_p Period_t + \beta_p Period_t \times log\_dist_i^j)}{\sum_{k=1}^{J} exp(\alpha_d^k + \gamma_t^k + \delta_p Period_t + \beta_p Period_t \times log\_dist_i^k)} \tag{3}$$

where subscript i represents the individuals in the city of Quito ($i = 1, \ldots, 2,544,382$), subscript d represents the district to which the individual belongs ($d = 1, \ldots, 51$), subscript t represents the daily time dimension ($t = 1, \ldots, 816$), subscript p represents the three temporal periods ($p = 1, \ldots, 3$), and subscript j represents the choice of visiting one of the four malls ($j = 0, \ldots, 4$, where 0 denotes the outside option, that of not visiting any of the four malls, which includes the choice of shopping online and not shopping at all). The main coefficients of interest are β_p, which capture the incremental effect of distance in periods 2 and 3. Note that in this specification the main effect of travel cost, i.e., the travel cost in period 1, is not estimated since it is collinear with district fixed effects. With α_d^j, we allow for district fixed effects to vary by mall (e.g., the product offering of certain malls might have better fit with individuals of certain districts). With γ_t^j, we allow for time effects to vary by mall (e.g., some seasonal promotions might be different in each mall). We operationalize the time effects as month of the year and day of the week fixed effects.[4] We consider that the main effect of period is the same across malls, δ_p, i.e., the average shock on propensity to visit (results are robust to a specification with different effects per mall, see Online Appendix EC.3.). We estimate the models with maximum likelihood.

Following our analytical model, we make three assumptions implicit in the logit framework (Train 2002). First, we assume that the decision of visiting a mall is independent across individuals and time, after controlling for district and time effects. We believe this assumption is reasonable to capture the average effect across the whole population, especially given the large sample size we employ for our estimation. We recognize that at individual level there might be autocorrelation in the decisions. For example, negative autocorrelation in the short-term and positive in the long-term, because a consumer that has visited the mall today might be less likely to visit the day immediately after but more likely to visit at some point in the future. We further assume that an individual visits only one of the four malls a day and that the independence from irrelevant alternatives property (IIA) holds in our setting. We empirically examine the plausibility of these last two assumptions.

Column (1) of Table 1 presents the estimation results of Equation (3), which support our predictions. The increase in travel cost due to COVID-19 affects shopping choices differently depending on the distance to destination, both for period 2 and 3. The negative estimates of β_p (-0.505 and -0.056, both $p < 0.001$) indicate that the farther the shopping destination, the larger the decrease in visit likelihood during post-COVID-19. The negative estimates of δ_p (-0.459 and -0.176, both $p < 0.001$) indicate that on average consumers reduced their likelihood to visit the malls after COVID-19. Both the main shock and the travel cost effects are stronger during lockdown. Importantly, taken these results together, the ratio of the base level and the interaction coefficients is

[4] Results are robust to alternative time fixed effects operationalizations (Online Appendix EC.2.)

smaller for period 3 than period 2 (-0.176/-0.056 < -0.459/-0.505) suggesting that the threshold from which shorter distances increases the likelihood to visit the mall is larger for period 2. That is, the higher the shock in travel cost the bigger the circle around the mall in which the likelihood to visit increases rather than decreases. In sum, with higher travel cost is more difficult to attract customers from far but is easier to capture and retain customer that are near.

Columns (2) to (6) of Table 1 introduce some variations in terms of variables and level of data aggregation. Although the model fit is reduced when district fixed effects are removed (LL, Pseudo R2, AIC, and BIC), the coefficients are consistent not only in terms of direction and significance but also in magnitude. In column (2), we remove the district fixed effects to estimate the main effect of distance. The magnitude of the distance coefficient (-1.380, $p < 0.001$) indicates that before COVID-19 for every unit increase in *log_dist* the propensity to visit decreased by a factor of 0.25 ($e^{-1.38}$). That is, every 2.63 km away from the mall three fourths of the customers are lost. The combination of the distance coefficient and its interactions with periods 2 and 3 indicate that for every unit increase in *log_dist* the propensity to visit decreased by a factor of 0.15 during lockdown and by 0.24 in the reopening.

Table 1. Main Results on the Effect of Distance on Shopping Destination Choice

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Period=2	-0.459***	-0.485***	-1.005***	-0.718***	-0.740***	-1.453***
	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)	(0.005)
Period=3	-0.176***	-0.180***	-0.462***	-0.266***	-0.267***	-0.643***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Period=2 x *log_dist*	-0.505***	-0.482***	-0.287***	-0.350***	-0.326***	-0.059***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Period=3 x *log_dist*	-0.056***	-0.053***	0.051***	-0.020***	-0.017***	0.124***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)

log_dist		-1.380*** (0.000)	-1.027*** (0.000)		-1.238*** (0.000)	-0.955*** (0.000)
Closest			0.995*** (0.001)			0.846*** (0.001)
Period=2 x *Closest*			0.460*** (0.002)			0.633*** (0.004)
Period=3 x *Closest*			0.256*** (0.001)			0.346*** (0.001)
Month FE	Included	Included	Included	Included	Included	Included
Day of the Week FE	Included	Included	Included	Not Included	Not Included	Not Included
District FE	Included	Not Included	Not Included	Included	Not Included	Not Included
Observations	2,076,215,712	2,076,215,712	2,076,215,712	290,059,548	290,059,548	290,059,548
Log-likelihood	-319,849,395	-334,389,672	-331,111,694	-119,352,202	-125,623,366	-124,431,021
Pseudo R2	0.158	0.119	0.128	0.163	0.119	0.127
AIC	639,699,343	668,779,498	662,223,548	238,704,908	251,246,838	248,862,155
BIC	639,704,712	668,780,996	662,225,105	238,709,315	251,247,765	248,863,134

Notes. Columns (1) to (3) are estimated with daily data. Columns (4) to (6) are estimated with weekday data. FE refers to fixed effects. In parentheses, standard errors are clustered at the district level.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Column (3) of Table 1 introduces a variable that accounts for whether the visited mall is the closest to the consumer. The larger magnitude of the negative estimates of δ_p compared to column (1) (-1.005 vs. -0.462, both $p < 0.001$) together with the positive interactions between having the mall as the closest and periods 2 and 3 (0.460 and 0.256, respectively, both $p < 0.001$) show that the negative shock on visit likelihood when travel cost increases is attenuated for the closest mall. This additional variable indirectly captures the effect of proximity to the shopping mall, and hence the effect of distance is weakened in all periods 1, 2 and 3 (-1.027 vs. -1.380, -0.287 vs. -0.482, and 0.051 vs. -0.053, respectively, all $p < 0.001$).

Columns (4) to (6) of Table 1 are estimated with aggregated weekly data during weekdays, i.e., unique weekly visitors from Monday to Friday (we do not have this aggregated data for weekends). We note some differences compared to the daily data estimation which suggest that visit frequency is also affected by the shock on travel cost. While the daily visit estimates capture the effect on visits (since daily visitors and daily visits are the same), the 5-day visit estimates capture the effect on the number of visitors during the week, thus omitting the frequency of visit of these visitors. The main effect due to COVID-19 is stronger with the 5-day aggregation than with daily data (-0.718 vs. -0.459, and -0.266 vs. -0.176, respectively, all $p < 0.001$). Hence, for customers nearby a shopping mall, the increase in travel cost increased the number of visits more than the number of visitors, suggesting that frequency increased. In contrast, the effect of distance is attenuated for 5-day aggregation compared to daily data (-0.350 vs. -0.505, and -0.020 vs. -0.056, respectively, all $p < 0.001$). Hence, for customers living far from the shopping mall, the increase in travel cost reduced the number of visits more than the number of visitors, suggesting that frequency decreased. Putting this in another way, the daily and weekday predicted probabilities intersect at $log\_dist_i^j = 1.7$ or 5.5 km. So that visits during the lockdown for customers at a distance of 5.5 km were multiplied by a factor of $e^{-0.459-0.505\times1.7} \approx 0.27$ and visitors were multiplied by the same factor $e^{-0.718-0.350\times1.7} \approx 0.27$. As a result, for distances larger than 5.5 km, frequency decreased, while it increased for shorter distances. Similarly, during the reopening period, frequency increased for distances shorter than 12.2 km, and decreased otherwise. Online Appendix EC.4. presents the factor change in predicted probabilities to visit with 5-day data, which is flatter than with daily data.

To ease the interpretation of the results, we plot in Figure 6 the factor change in predicted probabilities to visit a mall in periods 2 and 3 vs. period 1 of column (2) of Table 1. The combination of a higher cost of distance and a main decrease in visit likelihood during lockdown compared to the reopening phase can be noticed in that the factor change in predicted probability to visit for period 2 intersects the 100-line at the distance of around 500 m, while for period 3 intersect at 100 m. These intersections mark the thresholds from which the shock on travel cost provoked visits to increase for shorter distances and visits to decrease for longer distances. As distance increases, the effect of distance dominates over the main shock effect and the predictions of the different periods converge, see Online Appendix EC.5. for the predicted probabilities to visit instead of the factor change. At a distance of 1 km, the predicted probability to visit decreases from 8.5% in period 1 to 6.1% in period 2 and to 7.5% in period 3. At 2 km, the predicted probability to visit decreases to half in period 2 from 4.2% to 2.1% and in period 3 to 3.5%. At 10 km, the predicted probability to visit decreases six-fold in period 2 from 0.6% to 0.1%.

To explore further the relationship between travel cost and probability to visit the mall, we allow for a non-parametric functional form for distance coding the variable as categorical in intervals of deciles and quintiles. See estimates on Table 2. Overall, the estimates are consistent with the main results. Importantly, the effect of distance in period 1 is monotonically increasing in the whole range, except on the last decile (from -0.702 to -3.206 in the first and last deciles, respectively, both $p < 0.001$). However, although the shock on travel cost is not monotonic, as shown by the non-monotonicity of the interaction effects, the effect of distance in periods 2 and 3 remains largely monotonic (except the last deciles in both periods and the eight decile in period 3), as revealed by the sum of the main effects of distance and the interaction effects.

Table 2. Effect of Distance in Intervals

Variables	(1)	(2)	(3)	(4)
Period=2	-0.857*** (0.001)	-0.817*** (0.001)	-0.935*** (0.001)	-0.903*** (0.001)
Period=3	-0.236*** (0.000)	-0.227*** (0.000)	-0.242*** (0.000)	-0.231*** (0.000)
Period=2**Dist*=2	-0.385*** (0.002)	-0.445*** (0.002)		
Period=2**Dist*=3	-0.303*** (0.002)	-0.353*** (0.002)	-0.407*** (0.002)	-0.454*** (0.002)

$Period$=2*$Dist$=4	-0.808***	-0.869***		
	(0.003)	(0.003)		
$Period$=2*$Dist$=5	-1.191***	-1.229***	-1.084***	-1.103***
	(0.004)	(0.004)	(0.003)	(0.003)
$Period$=2*$Dist$=6	-1.079***	-1.107***		
	(0.006)	(0.006)		
$Period$=2*$Dist$=7	-0.963***	-0.982***	-0.915***	-0.929***
	(0.008)	(0.008)	(0.005)	(0.005)
$Period$=2*$Dist$=8	-0.986***	-1.027***		
	(0.008)	(0.008)		
$Period$=2*$Dist$=9	-0.839**	-0.870***	-0.831***	-0.837***
	(0.011)	(0.011)	(0.007)	(0.007)
$Period$=2*$Dist$=10	-0.955***	-0.965***		
	(0.009)	(0.009)		
$Period$=3*$Dist$=2	-0.017***	-0.023***		
	(0.001)	(0.001)		
$Period$=3*$Dist$=3	-0.029***	-0.038***	-0.014***	-0.020***
	(0.001)	(0.001)	(0.001)	(0.001)
$Period$=3*$Dist$=4	-0.004***	-0.007***		
	(0.001)	(0.001)		
$Period$=3*$Dist$=5	-0.175***	-0.185***	-0.155***	-0.164***
	(0.001)	(0.001)	(0.001)	(0.001)
$Period$=3*$Dist$=6	-0.119***	-0.125***		
	(0.002)	(0.002)		
$Period$=3*$Dist$=7	-0.135***	-0.143***	0.008***	0.001
	(0.002)	(0.002)	(0.001)	(0.001)
$Period$=3*$Dist$=8	0.101***	0.092***		
	(0.002)	(0.002)		
$Period$=3*$Dist$=9	0.140***	0.128***	0.023***	0.011***
	(0.003)	(0.003)	(0.002)	(0.002)
$Period$=3*$Dist$=10	-0.060***	-0.070***		
	(0.002)	(0.002)		
$Dist$=2		-0.702***		
		(0.000)		
$Dist$=3		-1.317***		-1.176***
		(0.001)		(0.000)
$Dist$=4		-1.409***		
		(0.001)		
$Dist$=5		-1.524***		-1.608***
		(0.001)		(0.001)
$Dist$=6		-2.322***		
		(0.001)		
$Dist$=7		-2.897***		-2.596***
		(0.001)		(0.001)
$Dist$=8		-2.949***		
		(0.001)		
$Dist$=9		-3.685***		-3.142***
		(0.002)		(0.001)
$Dist$=10		-3.206***		
		(0.001)		
Month FE	Included	Included	Included	Included
Day of the Week FE	Included	Included	Included	Not Included
District FE	Included	Not Included	Included	Not Included

Observations	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712
Log-likelihood	-319,882,036	-346,260,696	-319,909,719	-348,053,458
Pseudo R2	0.158	0.088	0.158	0.084
AIC	639,764,655	692,521,593	639,820,002	696,107,040
BIC	639,770,336	692,523,558	639,825,488	696,108,246

Notes. For columns (1) and (2), the variable *dist* (distance) is a categorical variable grouped in deciles. For columns (3) and (4), the variable *dist* is a categorical variable grouped in quintiles. *p<0.05, **p<0.01, ***p<0.001.

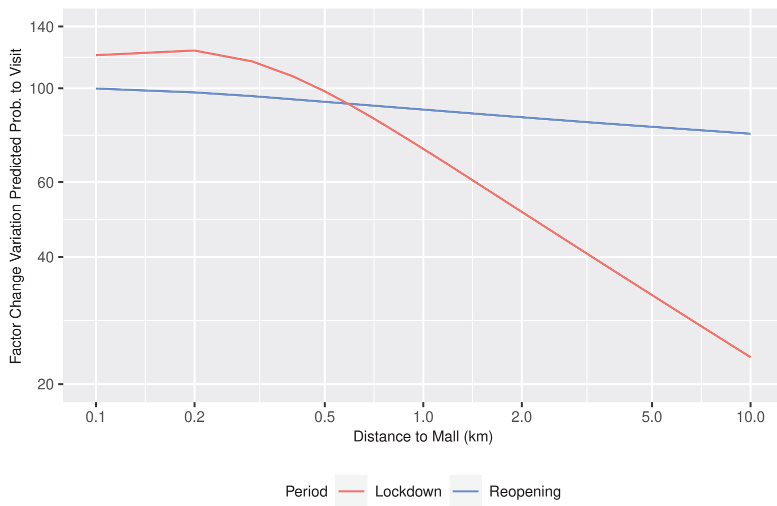
4.4. Heterogeneity Analysis

In this section we delve deeper into whether the effect of travel cost is driven by consumer choice. We posit that the higher the freedom to choose, the more relevant the travel cost faced by the consumer. We examine this proposition comparing whether distance has a differential effect on choices along three different factors related to ability of choice. First, we compare consumer choices during the week and weekend. We expect travel cost to be higher during the weekend since consumers have more alternatives to choose from during their free time and hence travel costs will be comparably more relevant in the choice process.[5] Second, we compare consumer choices on days with and without rain. Similarly, we expect travel cost to be more relevant in days without rain, given that during rainy days there are fewer acceptable shopping options to choose from (Martínez-de Albéniz and Belkaid 2021). Third, we examine the effect of social class on consumer choices. We expect wealthy consumers to experience higher travel cost since economic well-being provides more ability to choose (Sen 2000). Of the three heterogeneity analyses, weekend and rain dimensions vary with time while social class varies across districts. We augment Equation (3) to include the weekend dummy (variable WE) as follows (we do similarly to include rain and social class effects):

$$P_{it}^{j} := Pr(Y_{it}^{j} = j) = \frac{exp(\phi_{it}^{j})}{\sum_{k=1}^{J} exp(\phi_{it}^{k})} \tag{4}$$

[5] The results are robust to the inclusion of national holidays.

Figure 6 Effect of Distance on Shopping Destination Choice (MNL Prediction)



Notes. Each line represents the factor change variation in the predicted probability to visit with respect to period 1. Factor change variation above 100 means an increase in the predicted probability to visit and a decrease otherwise.

$$where \quad \phi_{it}^{j} = \alpha_d^j + \gamma_t^j + \delta_p Period_t + \beta_p Period_t \times log\_dist_i^j$$

$$+ \eta_p WE_t \times Period_t + \theta_p WE_t \times log\_dist_i^j + \lambda_p WE_t \times log\_dist_i^j \times Period_t$$

Column (1) of Table 3 presents the estimation results of Equation (4) and column (2) removes the district effects to allow the estimation of the main effects. Columns (3) to (6) mirror the first two columns for rain and social class. The results support our predictions: cost of travel is higher during weekends and rainy days and for wealthier people. To ease the interpretation of the results, we plot the predicted probabilities of columns (2), (4), and (6) in Figure 7. Figure 7 shows that distance has a larger effect on probability to visit on weekends, days without rain, and wealthy consumers.

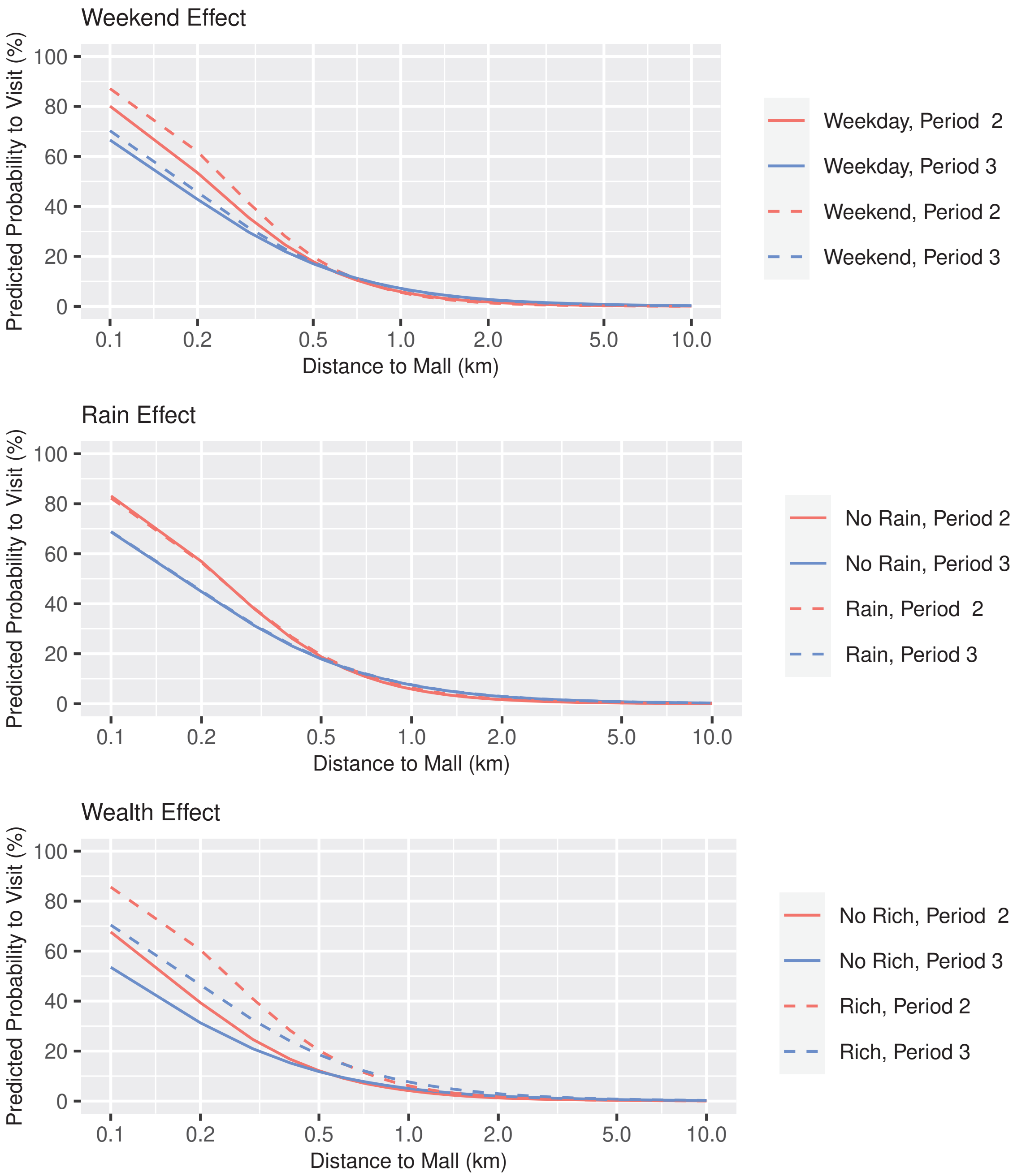
Table 3. Results on the Heterogeneity of the Effect of Distance

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Period=2	-0.476***	-0.496***	-0.404***	-0.425***	-0.276***	-0.412***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Period=3	-0.205***	-0.207***	-0.156***	-0.160***	-0.148***	-0.155***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	-0.475***	-0.456***	-0.473***	-0.454***	-0.481***	-0.379***

*Period=2***log_dist*	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
*Period=3***log_dist*	-0.059*** (0.000)	-0.0556*** (0.000)	-0.050*** (0.000)	-0.047*** (0.000)	-0.036*** (0.001)	-0.0298*** (0.000)
log_dist		-1.355*** (0.000)		-1.383*** (0.000)		-1.308*** (0.000)
WE		-0.126*** (0.001)				
*Period=2***WE*	0.094*** (0.003)	0.064*** (0.003)				
*Period=3***WE*	0.101*** (0.001)	0.101*** (0.001)				
*WE***log_dist*	-0.121*** (0.000)	-0.102*** (0.000)				
*Period=2***WE***log_dist*	-0.168*** (0.002)	-0.144*** (0.002)				
*Period=3***WE***log_dist*	0.019*** (0.001)	0.0192*** (0.001)				
No_rain			0.012*** (0.001)	0.013*** (0.001)		
*Period=2***No_rain*			-0.111*** (0.003)	-0.122*** (0.002)		
*Period=3***No_rain*			-0.041*** (0.001)	-0.041*** (0.001)		
*No_rain***log_dist*			0.006*** (0.000)	0.005*** (0.000)		
*Period=2***No_rain***log_dist*			-0.077*** (0.002)	-0.068*** (0.002)		
*Period=3***No_rain***log_dist*			-0.012*** (0.001)	-0.011*** (0.001)		
Rich						0.671*** (0.001)
*Period=2***Rich*					-0.315*** (0.003)	-0.132*** (0.003)
*Period=3***Rich*					-0.058*** (0.001)	-0.049*** (0.001)
*Rich***log_dist*						-0.083*** (0.000)
*Period=2***Rich***log_dist*					-0.052*** (0.002)	-0.187*** (0.002)
*Period=3***Rich***log_dist*					-0.028*** (0.001)	-0.035*** (0.001)
Month FE	Included	Included	Included	Included	Included	Included
Day of the Week FE	Included	Not Included	Included	Included	Included	Included
District FE	Included	Not Included	Included	Not Included	Included	Not Included
Observations	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712
Log-likelihood	-319,768,453	-334,690,407	-319,834,655	-334,375,061	-319,804,295	-333,223,059
Pseudo R2	0.158	0.119	0.158	0.120	0.158	0.123
AIC	639,537,468	669,380,931	639,669,875	668,750,287	639,609,151	666,446,284
BIC	639,542,935	669,382,079	639,675,361	668,751,902	639,614,598	666,447,899

*p<0.05, **p<0.01, ***p<0.001.

Figure 7 **Heterogeneity on the Effect of Distance on Shopping Destination Choice**



4.5. Robustness Checks

In this section we check the IIA assumption and the robustness of our results to a linear regression approach. IIA arises from the assumption that the error terms in Equation (3) are i.i.d. across malls. IIA may be violated if two or more of the malls are perceived by consumers as close substitutes. Hausman and McFadden (1984) demonstrate that if IIA is satisfied, then the estimated coefficients should be stable across choice sets. Therefore, we check the IIA assumption re-estimating the model with a subset of alternatives. Table 4 presents the estimates of Equation (3) excluding from the choice set one mall at a time, that is, reducing the choice set to three malls (hence, increasing the number of individuals that choose the outside option). We explore further the IIA assumption examining consumer choices of visiting a single mall with a binary logit (see Table 5). Examining each mall separately, the MNL collapses to a binary logit. In general, the results of both analysis are consistent in terms of magnitude and significance with the main results, suggesting that IIA is not violated in our setting.

Table 4. Check IIA Assumption: Exclude One Mall at a Time

Variables	excl. Mall A	excl. Mall B	excl. Mall C	excl. Mall D
Period=2	-0.458*** (0.001)	-0.340*** (0.002)	-0.590*** (0.002)	-0.444*** (0.001)
Period=3	-0.150*** (0.001)	-0.146*** (0.001)	-0.226*** (0.001)	-0.180*** (0.001)
Period=2 x *log_dist*	-0.450*** (0.001)	-0.515*** (0.001)	-0.517*** (0.001)	-0.538*** (0.001)
Period=3 x *log_dist*	-0.048*** (0.000)	-0.066*** (0.000)	-0.037*** (0.000)	-0.071*** (0.000)
Month FE	Included	Included	Included	Included
Day of the Week FE	Included	Included	Included	Included
District FE	Included	Included	Included	Included

Observations	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712
Log-likelihood	-245,576,970	-213,944,623	-219,109,714	-281,818,457
Pseudo R2	0.157	0.169	0.156	0.150
AIC	491,154,357	427,889,662	438,219,843	563,637,330
BIC	491,158,403	427,893,708	438,223,890	563,641,376

Notes. Column (1) excludes Mall A as alternative from the choice set, Column (2) excludes Mall B, Column (3) excludes Mall C, and Column (4) excludes Mall D.

*$p<0.05$, **$p<0.01$, ***$p<0.001$.

Table 5. Check IIA Assumption: Binary Logit

Variables	Mall A	Mall B	Mall C	Mall D	Mall A	Mall B	Mall C	Mall D
Period=2	-0.329***	-0.662***	-0.326***	-0.468***	-0.566***	-1.228***	-0.650***	0.941***
	(0.003)	(0.002)	(0.002)	(0.007)	(0.004)	(0.006)	(0.001)	(0.015)
Period=3	-0.222***	-0.219***	-0.112***	0.157***	-0.337***	-0.469***	-0.432***	0.441***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.005)
Period=2**log_dist*	-0.783***	-0.499***	-0.391***	-0.370***	-0.718***	-0.219***	-0.288***	-0.736***
	(0.002)	(0.002)	(0.001)	(0.004)	(0.003)	(0.003)	(0.002)	(0.005)
Period=3**log_dist*	-0.105***	-0.039***	-0.073***	-0.121***	-0.060***	0.079***	0.015***	-0.186***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
log_dist					-1.270***	-1.231***	-0.830***	-0.465***
					(0.000)	(0.001)	(0.000)	(0.001)
Closest					0.637***	0.363***	1.036***	2.728***
					(0.001)	(0.001)	(0.001)	(0.002)
*Closest***Period*=2					0.393***	0.636***	0.241***	-0.990***
					(0.003)	(0.001)	(0.005)	(0.009)
*Closest***Period*=3					0.170***	0.327***	0.253***	-0.213***
					(0.001)	(0.002)	(0.002)	(0.002)
Month FE	Included	Included	Included	Included	Included	Included	Included	Included
Day of the Week FE	Included	Included	Included	Included	Included	Included	Included	Included
District FE	Included	Included	Included	Included	Not Included	Not Included	Not Included	Not Included
Observations	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712
Log-likelihood	-74,902,876	-106,665,247	-100,971,228	-38,136,495	-76,419,732	-108,747,288	-104,737,736	-40,077,356
Pseudo R2	0.157	0.130	0.162	0.213	0.140	0.113	0.131	0.173
AIC	149,805,897	213,330,638	201,942,600	76,273,133	152,839,515	217,494,627	209,475,525	80,154,764
BIC	149,807,297	213,332,038	201,944,001	76,274,534	152,840,021	217,495,133	209,476,031	80,155,269

*$p<0.05$, **$p<0.01$, ***$p<0.001$.

Finally, we assess the robustness of the results to a log-log linear regression specification. We examine the relationship between the log of the number of individuals that visit a mall from each district as a function of the log of distance and other variables. The unit of observation is district-day and the estimates can be interpreted directly as elasticities. We present the results in Table 6 for each mall. The effects of distance are typically not significant in the estimations with district fixed effects. However, without district fixed effects, both the main effects of distance and its interactions with period 2 are negative and significant. The magnitude of the effects suggests that for a 1% increase in the distance to the mall, the volume of visits reduces between 1.5% and 1.8% for period 1, depending on the mall, between 1.8% and 2.3% for period 2, and between 1.4% and 1.8% for period 3.

Table 6. Robustness Check: Linear Regression

Variables	Mall A	Mall B	Mall C	Mall D	Mall A	Mall B	Mall C	Mall D
Period=2	-1.161*** (0.399)	-1.042*** (0.348)	-0.329 (0.230)	-1.407 (0.812)	-1.287** (0.058)	-1.229*** (0.065)	-0.921*** (0.110)	-1.176*** (0.174)
Period=3	-0.573*** (0.139)	-0.489*** (0.128)	-0.148 (0.200)	-0.012 (0.418)	-0.580*** (0.035)	-0.584*** (0.037)	-0.638*** (0.064)	0.080 (0.084)
Period=2**log_dist*	-0.333 (0.190)	-0.306 (0.167)	-0.435*** (0.099)	-0.076 (0.306)	-0.281*** (0.026)	-0.231*** (0.030)	-0.239*** (0.039)	-0.152** (0.059)
Period=3**log_dist*	0.126 (0.066)	0.149* (0.063)	-0.036 (0.093)	0.079 (0.153)	0.131*** (0.017)	0.187*** (0.018)	0.126*** (0.024)	0.049 (0.030)
log_dist					-1.855*** (0.011)	-1.878*** (0.012)	-1.529*** (0.016)	-1.873*** (0.023)
Population					0.258*** (0.002)	0.267*** (0.002)	0.255*** (0.002)	0.231*** (0.002)
Closest					0.333*** (0.016)	-0.472 (0.026)	0.285*** (0.025)	1.680*** (0.034)
Period=2**Closest*					0.515*** (0.070)	0.833*** (0.067)	0.452*** (0.062)	-0.237* (0.105)
Period=3**Closest*					0.136*** (0.029)	0.428*** (0.039)	0.374*** (0.035)	-0.095* (0.045)
Month FE	Included	Included	Included	Included	Included	Included	Included	Included
Day of the Week FE	Included	Included	Included	Included	Included	Included	Included	Included
District FE	Not Included	Not Included	Not Included	Not Included	Not included	Not included	Not included	Not included
Observations	41,616	41,616	41,616	41,616	41,616	41,616	41,616	41,616
R2	0.531	0.532	0.389	0.371	0.710	0.706	0.750	0.599

Notes. Dependent variable measured in logs. *Population* measured in ten thousands inhabitants.

*$p<0.05$, **$p<0.01$, ***$p<0.001$.

5. Generalizability: Cities of Guayaquil, Manta, and Barcelona

In this section, we assess the generalizability of our empirical results to three additional cities in two countries. These are the cities of Guayaquil and Manta in Ecuador and the city of Barcelona in Spain. We obtained the same data as for the analysis of Quito but for a single mall in each city. For the city of Guayaquil we considered 33 districts with a population of 1,009,117, for Manta 17 districts with a population of 233,061, and for Barcelona 73 districts with a population of 1,621,481.[6] See Online Appendix EC.6. for a replication of the model-free evidence analysis for these cities.

Table 7 presents the estimates of Equation (3) for these three cities. Note that examining the choice of visiting a mall, the MNL collapses to a binary logit. Overall, the results are consistent with our prediction and empirical findings in Quito in terms of direction and significance of the estimates: (1) travel cost is negative, (2) the shock on travel cost is negative with larger magnitude in period 2 than period 3, (3) the main decrease on visit likelihood is larger in period 2 than period 3, (4) the higher the travel cost the bigger the circle around the mall in which the likelihood to visit increases (obtained from the ratio of the base level and the interaction coefficients), and (5) having the mall as closest increases the likelihood of visit, more so when travel cost increases. Focusing on the magnitude of the effects, we observe some interesting differences. From a country perspective, the magnitude of the shock on travel cost appears to be deeper and to last longer in Barcelona, Spain than in the three Ecuadorian cities (in column (3), the interactions of distance with periods 2 and 3 are -1.088 and -0.936, respectively, both $p < 0.001$). Also, Barcelona, Spain seems to present a V-shaped recovery since the main effect of the shock on visit likelihood is positive for period 3 (0.778, $p < 0.001$). Within Ecuador, the shock on travel cost during the lockdown is larger in Guayaquil and Manta than in Quito (the interaction of distance with periods 2 for Guayaquil and Manta is -0.666 and -0.908, respectively, both $p < 0.001$). However, the shock on

[6] Note that the lockdown periods are different in each city. In Guayaquil, the lockdown went from March 17 to May 20, 2020 (Guayaquil 2020). In Manta, the lockdown went from March 17 to June 10, 2020 (Gobierno de Manta 2020). In Barcelona, the lockdown went from March 15 to June 15, 2020 (BOE 2020).

period 3 is larger in Quito, especially compared to Guayaquil (the interaction of distance with periods 3 for Guayaquil and Manta is -0.002 and -0.048. Likewise, the main effect of the shock on visit likelihood is larger in Guayaquil and Manta than in Quito (periods 2 and 3 for Guayaquil and Manta are -0.857 and -0.381, and -0.555 and -0.278, respectively, all $p < 0.001$). See Online Appendix EC.7. for an estimation that allows for different shock effects on travel cost per month. See Online Appendix EC.8. for a robustness check with a log-log linear regression specification.

Finally, we note some heterogeneity in the magnitudes of the threshold around the mall in which visits increase due to a higher travel cost. During lockdown, the thresholds were 900 m, 800 m, and 400 m in Guayaquil, Manta, and Barcelona, respectively, compared to 500 m in Quito. During the reopening, the thresholds were 100 m, 100 m, and 2,100 m in Guayaquil, Manta, and Barcelona, respectively, compared to 100 m in Quito. Hence, in the three Ecuadorian cities, where the shock on travel cost is combined with a main decrease effect in both periods, the threshold is larger in the lockdown period than in the reopening. On the contrary, in Barcelona, Spain, the reopening period is characterized by a positive main increase effect (indication of a V-shaped recovery) which combined with the shock on travel cost results in a larger threshold in the reopening phase compared to the lockdown.

Table 7. Generalizability: Logit Estimates for the Cities of Guayaquil, Manta, and Barcelona

Variables	Mall Gua	Mall Man	Mall Bcn	Mall Gua	Mall Man	Mall Bcn
Period=2	-0.857*** (0.004)	-0.555*** (0.010)	-0.558*** (0.006)	-1.553*** (0.011)	-0.599*** (0.026)	-1.134*** (0.009)
Period=3	-0.381*** (0.001)	-0.278*** (0.003)	0.778*** (0.002)	-0.553*** (0.002)	-0.506*** (0.006)	0.171*** (0.003)
Period=2*log_dist	-0.666*** (0.003)	-0.908*** (0.008)	-1.088*** (0.005)	-0.312*** (0.006)	-0.885*** (0.017)	-0.682*** (0.006)
Period=3*log_dist	-0.002* (0.001)	-0.048*** (0.002)	-0.936*** (0.002)	0.084*** (0.001)	0.079*** (0.003)	-0.533*** (0.002)
log_dist				-0.507*** (0.001)	-0.852*** (0.002)	-1.164*** (0.001)
Closest				1.934*** (0.002)	0.490*** (0.003)	0.199*** (0.001)
*Closest***Period*=2				0.839*** (0.010)	0.079*** (0.016)	0.299*** (0.007)
*Closest***Period*=3				0.269*** (0.003)	0.216*** (0.004)	0.363*** (0.002)
Month FE	Included	Included	Included	Included	Included	Included
Day of the Week FE	Included	Included	Included	Included	Included	Included
District FE	Included	Included	Included	Not included	Not included	Not included
Observations	823,439,472	190,177,776	1,003,696,765	823,439,472	190,177,776	1,003,696,765
Log-likelihood	-68,049,315	-14,198,222	-42,744,156	-74,822,885	-14,793,789	-43,171,037
Pseudo R2	0.161	0.094	0.076	0.08	0.056	0.066
AIC	136,098,739	28,396,519	85,488,500	149,645,822	29,587,630	86,342,126
BIC	136,099,739	28,397,168	85,490,260	149,646,304	29,588,074	86,342,613

*p<0.05, **p<0.01, ***p<0.001.

6. Discussion

Although the literature posits that an increase in travel cost would reduce shopping visits, anecdotal evidence suggests that this might not always be true. To this end, we build on the gravity law to develop a new model of consumer shopping destination choices. Our model predicts that with an increase in travel cost consumers tend to substitute shopping visits at distant venues for nearby alternatives. We empirically examine our predictions in four cities from two countries with data on customer visits to seven shopping malls. We exploit the COVID-19 pandemic as a natural shock on travel cost to identify the relationship between cost of travel and mall visits. Our empirical results validate our analytical predictions in that in each geography there is a threshold from the shopping destination below which visits increased due to increased travel costs. In our empirical setting, this threshold is in the order of 500 meters during the lockdown period and 100 meters after the lockdown. Furthermore, we show that the travel cost effect is strengthened in higher ability to choose circumstances, that is, during weekends, in no-rain days, and for higher social class individuals.

Our new model and findings not only advances the theoretical understanding of the effect of travel cost on shopping choices, but has implications for managers and policy makers. Our findings informs retailers on what to expect when future changes on travel cost occur, not only in terms of overall customer traffic but also customer profile depending on their origin. Moreover, our new choice model helps quantify the area of influence around the retail location, hence, helping design a store networks by evaluating the appropriate store density. For policy makers, this research sheds light on the effects on local retailing of policies that affect mobility or the city landscape, for example, urban mobility restrictions or access to public transportation. In sum, this paper contributes academically to the marketing and customer choice modeling literature, and managerially to retailers and urban policy makers.

Limitations of our paper suggest useful directions for further research. First, due to privacy reasons and GDPR regulation, our mall visits data is aggregated at district level. Hence, we can make inferences at individual level on the evolution of visits and visitors, but not on visit frequency. Although our results suggest that frequency might be affected differently depending on customer location, further research could explore more in detail this phenomenon. Second, the availability of customer data aggregated at group level is pervasive and growing in the marketing field, given the regulatory trend of restricting the use of individual data (e.g., online cookies). Therefore, future research could investigate how to infer heterogeneous effects on visit frequency from aggregated data. Third, a limitation of our data is that we observe few data points at short distances from the malls. Future research with richer data on this aspect could explore more in detail the magnitude of the areas of influence around stores. Additionally, it would be interesting to compare how this area of influence changes across other geographies, and even identify factors that affect the size of the area of influence.

Acknowledgments

The Authors thank the Social Trends Institute for financial support to this research. The Authors thank Nerea Frias for superb research assistance.

References

Alexander D, Karger E (2020) Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior. Working Paper 2020-12, Federal Reserve Bank of Chicago, Chicago.

Alé-Chilet J, Atal JP, Domínguez P (2020) Activity and the incidence of eemergencies: Evidence from daily data at the onset of a pandemic Pier Working paper 20-016, PCPSE, Philadelphia, PA.

Anderson JE (2011) The gravity model. *Annu. Rev. Econ.* 3(1):133–160.

Aranda JL (2019) ¿Ha arruinado Madrid Central a las tiendas? *El Pais* (May 30), https://elpais.com/economia/2019/05/29/actualidad/1559154043_468446.html.

Beine MAR, Docquier F, Ozden C (2009) Diasporas Working Paper No. 4984, World Bank Policy Research.

Beiró MG, Bravo L, Caro D, Cattuto C, Ferres L, Graells-Garrido E (2018) Shopping mall attraction and social mixing at a city scale. *EPJ Data Sci.* 7(28).

Bell DR, Gallino S, Moreno A (2020) Customer supercharging in experience-centric channels. *Manag. Sci.* 66(9):4096–4107.

Bell DR, Ho TH, Tang CS (1998) Determining where to shop: fixed and variable costs of shopping. *J. Mark. Res.* 35(3):352–369.

Benvenuty L (2021) La supermanzana de sant antoni lastra al comercio. *La Vanguardia* (July 21), https://www.lavanguardia.com/local/barcelona/20210721/7614066/supermanzana-sant-antoni-lastra-comercio.html.

Blum BS, Goldfarb A (2006) Does the internet defy the law of gravity? *J. Int. Econ.* 70(2):384–405.

BOE (2020) Orden snd/520/2020, de 12 de junio. Accessed July 29, 2021, Boletín Oficial del Estado, Spain. Num. 166, Sec. I., 39946–39951, https://www.boe.es/boe/dias/2020/06/13/pdfs/BOE-A-2020-6088.pdf.

Bounie D, Camara Y, Galbraith J (2020) Consumers' mobility, expenditure and online-offline substitution response to covid-19: Evidence from french transaction data. Technical report, 02566443, HAL. Telecom, Paris.

Burtch G, Ghose A, Wattal S (2014) Cultural differences and geography as determinants of online prosocial lending. *MIS Q* 38(3):773–794.

Campos-Vazquez RM, Esquivel G (2021) Consumption and geographic mobility in pandemic times. evidence from mexico. *Rev. Econ. Househ.* 19:353–371.

Caro F, Kök AG, Martínez-de Albéniz V (2020) The future of retail operations. *Manuf. Serv. Oper. Manag.* 22(1):47–58.

Carvalho VM, Hansen S, Álvaro Ortiz, García JR, Rodrigo T, Mora SR, Ruiz J (2020) Tracking the covid-19 crisis with high-resolution transaction data. Technical report, CEPR discussion papers, 14642.

Chetty R, Friedman JN, Hendren N, Stepner M, Team TOI (2020) The economic impacts of covid-19: Evidence from a new public database built using private sector data Working paper 27431, NBER, Cambridge, MA.

Chintagunta PK, Chu J (2021) Geography as branding: Descriptive evidence from taobao. *Quant. Mark. Econ.* 19:53–92.

Chintagunta PK, Chu J, Cebollada J (2012) Quantifying transaction costs in online/off-line grocery channel choice. *Manag. Sci.* 31(1):96–114.

Chronopoulos DK, Lukas M, Wilson JOS (2020) Consumer spending responses to the COVID-19 pandemic: An assessment of Great Britain. Available at SSRN: https://ssrn.com/abstract=3586723.

Converse PD (1949) New laws of retail gravitation. *J. Mark.* 14(3):379–384.

Forman C, Ghose A, Goldfarb A (2009) Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Manage Sci.* 55(1):47–57.

Gautschi DA (1981) Specification of patronage models for retail center choice. *J. Mark. Res.* 18(2):162–174.

Ghosh A, Craig CS (1983) Formulating retail location strategy in a changing environment. *J. Mark.* 47(3):56–68.

Gobierno de Ecuador (2020) Desde el 1 de junio se cambian algunas reglas de la semaforización por la emergencia sanitaria. Accessed July 29, 2021, https://www.ministeriodegobierno.gob.ec/desde-el-1-de-junio-se-cambian-algunas-reglas-de-la-semaforizacion-por-la-emergencia-sanitaria/.

Gobierno de Manta (2020) Terminal terrestre de manta vuelve a su actividad bajo estricto control. Accessed July 29, 2021, https://manta.gob.ec/terminal-terrestre-de-manta-vuelve-a-su-actividad-bajo-estricto-control-reduccion-de-aforo-senalizacion-preventiva-e-instalacion-de-13-lavamanos/.

Gomez-Herrera E, Martens B, Turlea G (2014) The drivers and impediments for cross-border e-commerce in the EU. *Inf. Econ. Policy* 28:83–96.

Grogger J, Hanson GH (2011) Income maximization and the selection and sorting of international migrants. *J. Dev. Econ.* 95(1):42–57.

Guayaquil GM (2020) Solo los vehiculos con las placas terminadas en 6,7,8 y 0 pueden circular el jueves 28 de mayo. Accessed July 29, 2021, https://guayaquil.gob.ec/solo-los-vehiculos-con-las-placas-terminadas-en-6-7-8-y-0-pueden-circular-el-jueves-28-de-mayo/.

Hausman J, McFadden D (1984) Specification tests for the multinomial logit model. *Econometrica* 52(5):1219–1240.

He C, Wang T, Luo X, Luo Z, Guan J, Gao H, Zhu K, Feng L, Xu Y, Cheng Y, Hu YJ (2020) Surviving COVID-19: Recovery curves of mall traffic in China. Research Paper No. 3613294, Georgia Tech Scheller College of Business.

Hortaçsu A, Martínez-Jerez FA, Douglas J (2009) The geography of trade in online transactions: Evidence from eBay and MercadoLibre. *Am. Econ. J.: Microecon.* 1(1):53–74.

Hu W (2016) What New York can learn from Barcelona's 'superblocks'. *New York Times* (September 30), https://www.nytimes.com/2016/10/02/nyregion/what-new-york-can-learn-from-barcelonas-superblocks.html.

Huff DL (1964) Defining and estimating a trading area. *J. Mark.* 28(3):34–38.

Jia JS, Lu X, Yuan Y, Xu G, Jia J, Christakis NA (2020) Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* 582:389–394.

Jung WS, Wang F, Stanley HE (2008) Gravity model in the Korean highway. *EPL* 81(4):48005.

Kahn BE (2021) *The Shopping Revolution, Updated and Expanded Edition: How Retailers Succeed in an Era of Endless Disruption Accelerated by COVID-19* (Wharton Digital Press).

Krings G, Calabrese F, Ratti C, Blondel VD (2009) Urban gravity: A model for inter-city telecommunication flows. *J. Stat. Mech. Theory Exp.* L07003.

Lee ML, Pace RK (2005) Spatial distribution of retail sales. *J. Real Estate Finance Econ.* 31(1):53–69.

Lendle A, Olarreaga M, Schropp S, Vézina P (2016) There goes gravity: eBay and the death of distance. *Econ. J.* 126(591):406–441.

Li Y, Liu L (2012) Assessing the impact of retail location on store performance: A comparison of Wal-Mart and Kmart stores in Cincinnati. *Appl. Geogr.* 32(2):591–600.

Marshall G, Pires T (2018) Measuring the impact of travel costs on grocery shopping. *Econ. J.* 128(614):2538–2557.

Martínez-de Albéniz V, Belkaid A (2021) Here comes the sun: Fashion goods retailing under weather fluctuations. *Eur. J. Oper. Res.* 294(3):820–830.

Martínez-de-Albéniz V, Aparicio D, Balsach J (2021) The effects of covid-19 on the retail economy: Store restrictions, consumer panic, and travel shocks, Working paper, IESE Business School, Spain.

McFadden D (1974) Conditional logit analysis of qualitative choice behavior. Zarembka P, *Frontiers in Econometrics.* (Academic Press, New York), 105–142.

NASA POWER Project (2003) Prediction of Worldwide Energy Resource. Accessed July 29, 2021, https://power.larc.nasa.gov/data-access-viewer/.

Nazir S (2019) UK high street footfall falls 10% in last 7 years. *Retail Gazette* (October 14), https://www.retailgazette.co.uk/blog/2019/10/uk-high-street-footfall-falls-10-last-7-years/.

Okoruwa AA, Terza JV, Nourse HO (1988) Estimating patronization shares for urban retail centers: An extension of the poisson gravity model. *J. Urban. Econ.* 24(3):241–259.

Reilly WJ (1931) *The Law of Retail Gravitation* (W.J.Reilly, New York).

Relihan L, Ward M, Wheat CW, Farrell D (2020) The early impact of covid-19 on local commerce: changes in spend across neighborhoods and online. Report, LSE Research Online Documents on Economics 105060, London School of Economics and Political Science.

Rust RT, Donthu N (1995) Capturing geographically localized misspecification error in retail store choice models. *J. Mark. Res.* 32(1):103–110.

Sen A (2000) *Development as Freedom* (Alfred A. Knopf Inc, New York).

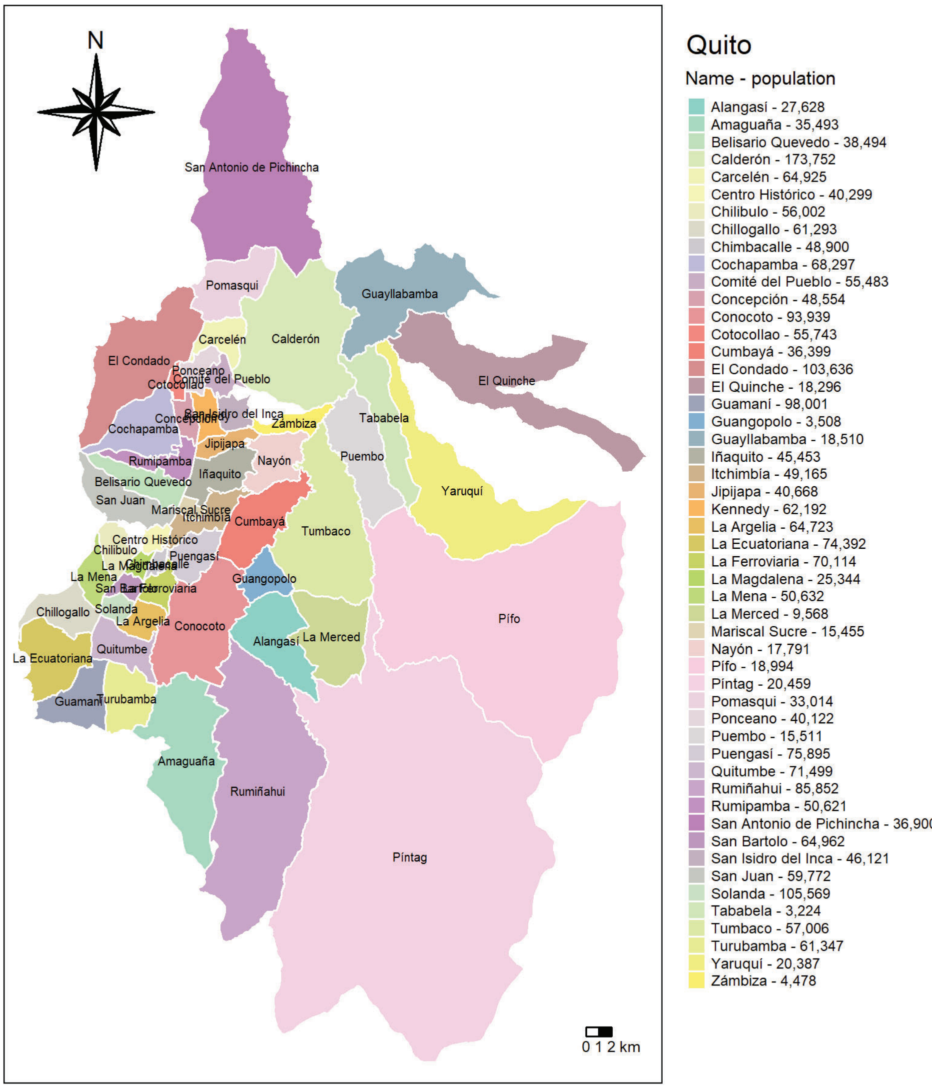
Simini F, González MC, Maritan A, Barabási AL (2012) A universal model for mobility and migration patterns. *Nature* 484:96–100.

Train K (2002) *Discrete Choice Methods with Simulation* (Cambridge University Press, New York).

Online Appendix

EC.1. Quito City Map and Descriptives

Figure EC.1.1 51 Districts Included in the Analysis.



EC.2. Robustness Checks: Alternative Variable Measurements

We assess the robustness of the results to alternative variable measurements. Table EC.2.1 considers four periods (instead of three as in the main analysis) which takes into account a minor change in the legal restrictions on the proportion of retail capacity allowed to be occupied. We consider the measurement of the variable distance in levels (instead of in logs as in the main analysis). We consider two measurements for the month fixed effects: one dummy per month (27 dummies) and one dummy per calendar month (12 dummies). Finally, Table EC.2.2 considers a period measurement with a level for each month since the start of COVID-19.

Table EC.2.1. Robustness Check: Distance Effect in 4 Periods

Variables	(1)	(2)	(3)	(4)
Period=2	-0.458***	-0.932***	-0.671***	-1.161***
	(0.001)	(0.002)	(0.001)	(0.002)
Period=3	-0.281***	-1.282***	-0.346***	-1.350***
	(0.001)	(0.004)	(0.001)	(0.003)
Period=4	-0.136***	-1.321***	-0.180***	-1.371***
	(0.001)	(0.004)	(0.004)	(0.004)
Period=2**dist*	-0.506***	-0.510***	-0.086***	-0.083***
	(0.001)	(0.001)	(0.000)	(0.000)
Period=3**dist*	-0.079***	-0.088***	-0.009***	-0.009***
	(0.000)	(0.001)	(0.000)	(0.000)
Period=4**dist*	-0.048***	-0.061***	-0.004***	-0.005***
	(0.000)	(0.000)	(0.000)	(0.000)
Month FE	Included	Included	Included	Included
Day of the Week FE	Included	Included	Included	Included
Area FE	Included	Included	Included	Included
Observations	2,076,215,712	2,076,215,712	2,076,215,712	2,076,215,712
Log-likelihood	-319,796,334	-319,434,680	-319,839,369	-319,484,750
Pseudo R2	0.157	0.158	0.157	0.158
AIC	639,593,224	638,870,035	639,679,294	638,970,176
BIC	639,598,632	638,876,611	639,684,702	638,976,752

Notes. The variable *dist* is measured in logs (columns 1 and 2), in levels (km) in 3 and 4. Columns 1 and 3 include *Month* as a categorical with periods 1-12 (Jan-Dec). Columns 2 and 4 include *Month* with the database monthly periods (1-27).
*p<0.05, **p<0.01, ***p<0.001.

Table EC.2.2. Robustness Check: Distance Effect by Month

Variables	(1)	(2)
Period=2	-0.621*** (0.003)	-0.673*** (0.003)
Period=3	-0.410*** (0.002)	-0.440*** (0.002)
Period=4	-0.289*** (0.002)	-0.304*** (0.002)
Period=5	-0.600*** (0.002)	-0.598*** (0.002)
Period=6	-0.197*** (0.002)	-0.194*** (0.002)
Period=7	-0.056*** (0.002)	-0.058*** (0.002)
Period=8	-0.331*** (0.002)	-0.338*** (0.002)
Period=9	-0.211*** (0.002)	-0.213*** (0.002)
Period=10	-0.287*** (0.002)	-0.290*** (0.002)
Period=11	-0.219*** (0.001)	-0.217*** (0.001)
Period=12	-0.064*** (0.001)	-0.069*** (0.001)
Period=13	-0.032*** (0.001)	-0.044*** (0.001)
Period=14	0.032*** (0.001)	0.022*** (0.001)
Period=2 x *log_dist*	-0.721*** (0.002)	-0.675*** (0.002)
Period=3 x *log_dist*	-0.629*** (0.002)	-0.602*** (0.001)
Period=4 x *log_dist*	-0.402*** (0.001)	-0.388*** (0.001)
Period=5 x *log_dist*	-0.107*** (0.001)	-0.105*** (0.001)
Period=6 x *log_dist*	-0.077*** (0.001)	-0.078*** (0.001)
Period=7 x *log_dist*	-0.076*** (0.001)	-0.075*** (0.001)
Period=8 x *log_dist*	-0.046*** (0.001)	-0.040*** (0.001)
Period=9 x *log_dist*	-0.047*** (0.001)	-0.045*** (0.001)
Period=10 x *log_dist*	-0.062*** (0.001)	-0.058*** (0.001)
Period=11 x *log_dist*	-0.022*** (0.001)	-0.023*** (0.001)
Period=12 x *log_dist*	-0.051*** (0.001)	-0.047*** (0.001)
Period=13 x *log_dist*	-0.059*** (0.001)	-0.051*** (0.001)

Period=14 x *log_dist*	-0.050*** (0.001)	-0.043*** (0.001)
log_dist		-1.380*** (0.000)
Month FE	Included	Included
Day of the Week FE	Included	Included
District FE	Included	Not Included
Observations	2,076,215,712	2,076,215,712
Log-likelihood	-319,551,175	-334,093,852
Pseudo R2	0.159	0.120
AIC	639,102,947	668,187,903
BIC	639,108,744	668,189,829

Note. Model 1 and 2 are estimated using a new variable for period defined as; *Period*=1 ("Pre-covid"), *Period*=2 (last half of March 2020), *Period*=3 for the month of April 2020 and so on.
*$p<0.05$, **$p<0.01$, ***$p<0.001$.

EC.3. Robustness Checks: Different Period Effect per Mall

Table EC.3.1. Different Main Effect by Mall

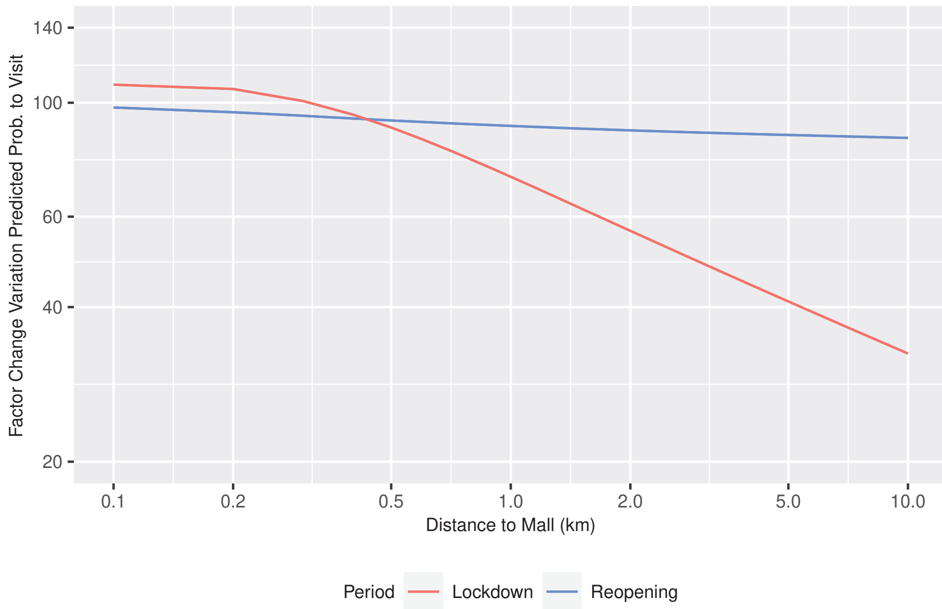
Variables	(1)
Period=2_*Mall A*	-0.646*** (0.002)
Period=3_*Mall A*	-0.285*** (0.001)
Period=2_*Mall B*	-0.672*** (0.002)
Period=3_*Mall B*	-0.182*** (0.001)
Period=2_*Mall C*	-0.209*** (0.002)
Period=3_*Mall C*	-0.121*** (0.001)
Period=2_*Mall D*	-0.222*** (0.003)
Period=3_*Mall D*	0.045*** (0.001)
Period=2**log_dist*	-0.510*** (0.001)
Period=3**log_dist*	-0.068*** (0.000)
Month FE	Included
Day of the Week FE	Included
District FE	Included
Observations	2,076,215,712
Log-likelihood	-319,752,049
Pseudo R2	0.158
AIC	639,504,662
BIC	639,510,148

Note. Variable *Period* different effect for each alternative.

*p<0.05, **p<0.01, ***p<0.001.

EC.4. Distance effect

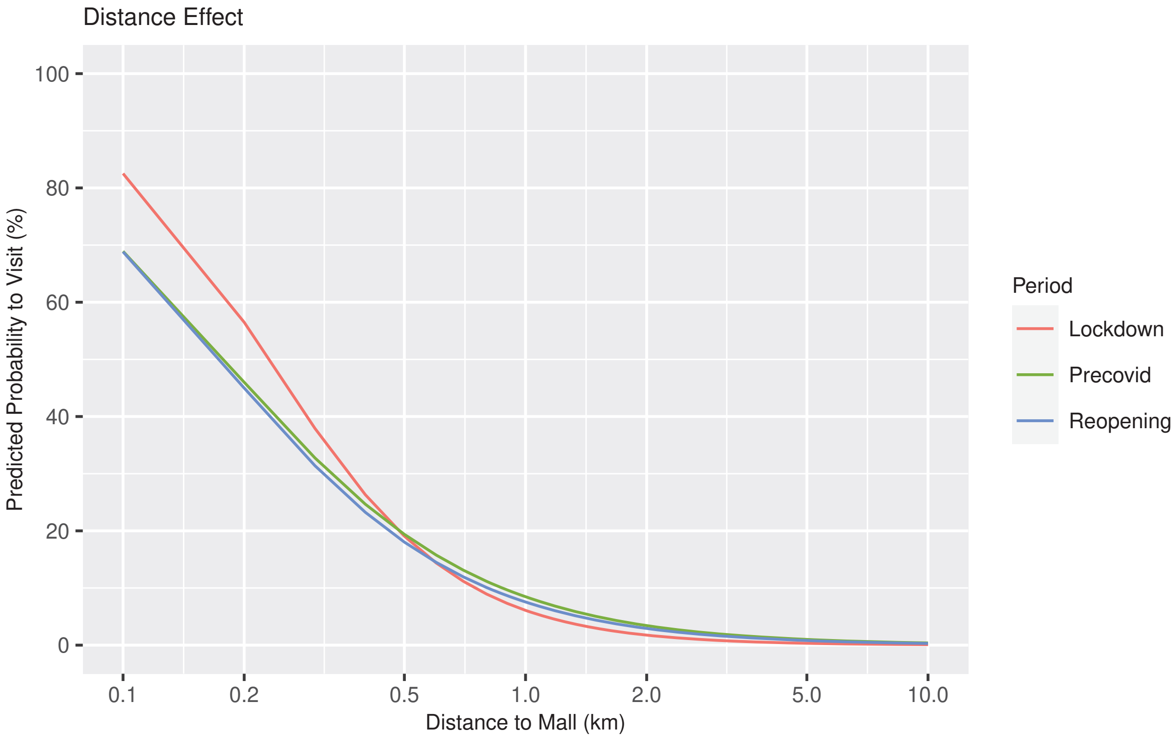
Figure EC.4.1 Effect of Distance on Shopping Destination Choice (Week aggregation)



Notes. Each line represents the factor change variation in the predicted probability to visit with respect to period 1. Factor change variation above 100 means an increase in the predicted probability to visit and a decrease otherwise.

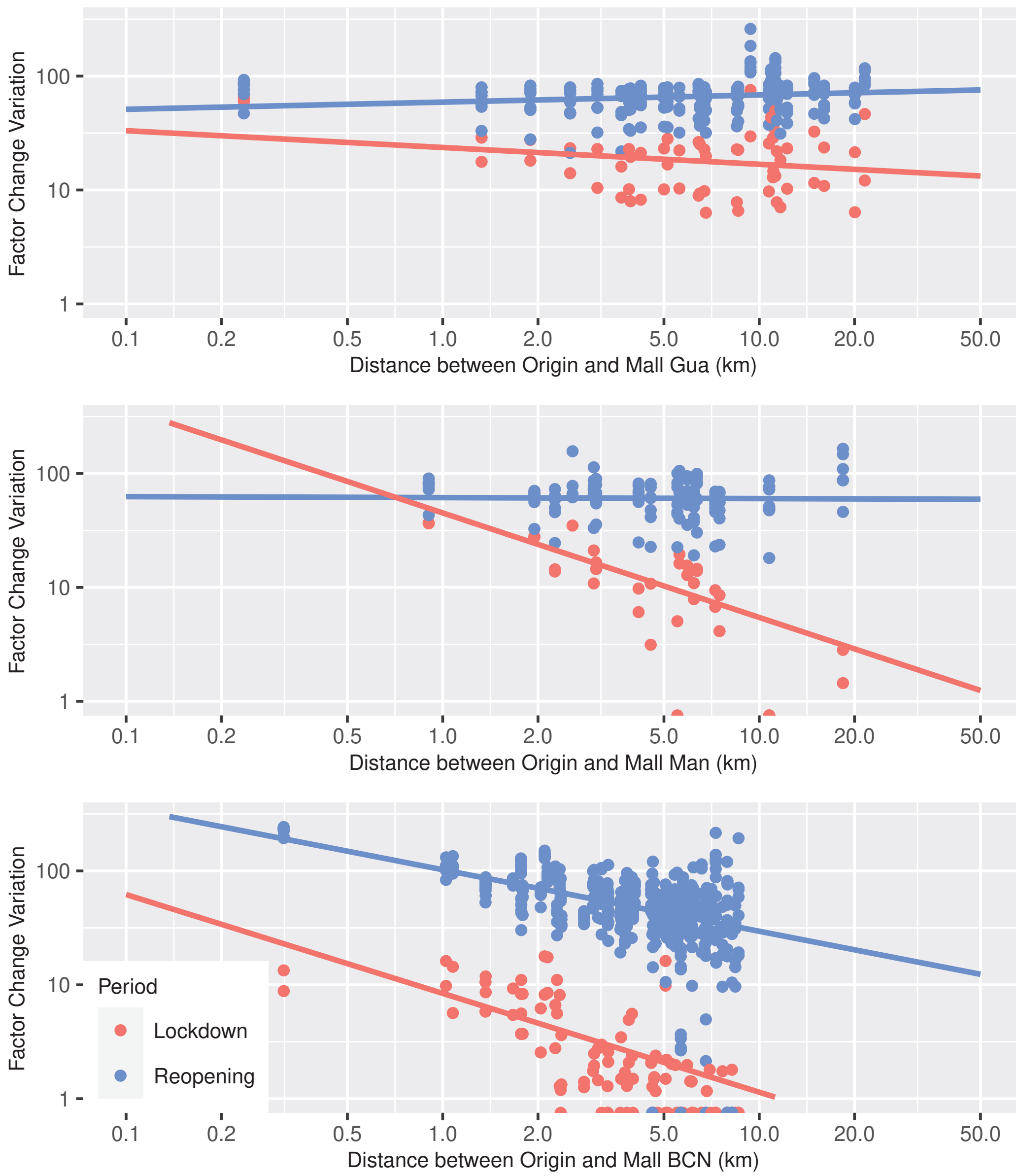
EC.5. Predicted Probabilities to Visit

Figure EC.5.1 Effect of Distance on Shopping Destination Choice



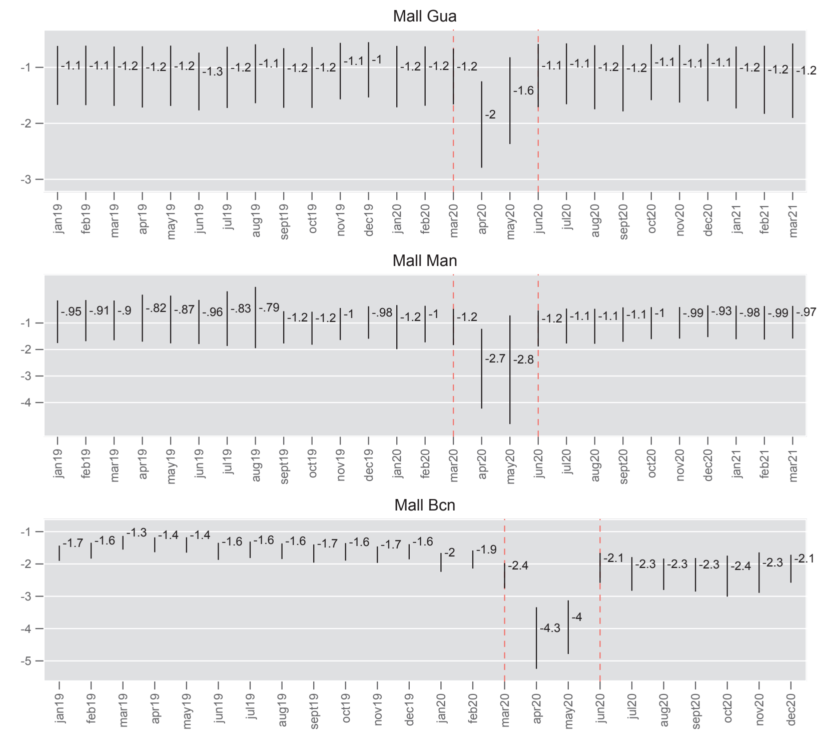
EC.6. Model-free Evidence: Cities of Guayaquil, Manta, and Barcelona

Figure EC.6.1 Factor Change in the Number of Daily Visits During Lockdown (Period 2) and Reopening (Period 3).



Notes. Factor change variation represents visits post-COVID-19 divided by pre-COVID-19, A number above 100 indicates an increase in the number of visits and a decrease otherwise. The two lines are the linear OLS predictions of the relationship between distance and change in visits.

Figure EC.6.2 Estimates of the Relationship between Distance and Visits (log-log Regression) for 3 Different Malls.



Notes. Coefficient estimates of *log_dist* for Equation (1). Mall Gua and Mall Man include the months of Jan '21 - Mar '21.

EC.7. Travel Cost by Month in Guayaquil, Manta and Barcelona

Table EC.7.1. Robustness Check: Distance Effect by Month for the 3 Different Cities

Variables	Mall Gua	Mall Man	Mall BCN	Mall Gua	Mall Man	Mall BCN
Period=2	-0.797*** (0.008)	-0.442*** (0.024)	-1.949*** (0.027)	0.077*** (0.010)	-0.113*** (0.028)	-2.218*** (0.023)
Period=3	-0.885*** (0.006)	-0.463*** (0.017)	-1.828*** (0.016)	-0.046*** (0.007)	-0.176*** (0.019)	-2.103*** (0.014)
Period=4	-0.789*** (0.005)	-0.498*** (0.016)	-1.405*** (0.012)	-0.311*** (0.007)	-0.253*** (0.018)	-1.632*** (0.010)
Period=5	-1.124*** (0.004)	-0.870*** (0.011)	0.868*** (0.005)	-0.965*** (0.006)	-0.791*** (0.013)	0.692*** (0.004)
Period=6	-0.623*** (0.004)	-0.251*** (0.010)	0.772*** (0.005)	-0.531*** (0.005)	-0.205*** (0.012)	0.608*** (0.004)
Period=7	-0.549*** (0.004)	-0.160*** (0.010)	0.645*** (0.006)	-0.431*** (0.005)	-0.116*** (0.012)	0.482*** (0.005)
Period=8	-0.367*** (0.004)	-0.349*** (0.009)	0.758*** (0.005)	-0.285*** (0.005)	-0.319*** (0.011)	0.581*** (0.004)
Period=9	-0.260*** (0.004)	-0.229*** (0.009)	0.809*** (0.005)	-0.186*** (0.005)	-0.212*** (0.010)	0.628*** (0.004)
Period=10	-0.321*** (0.004)	-0.359*** (0.009)	0.734*** (0.005)	-0.235*** (0.005)	-0.334*** (0.010)	0.535*** (0.004)
Period=11	-0.272*** (0.003)	-0.361*** (0.008)	0.971*** (0.004)	-0.246*** (0.005)	-0.360*** (0.009)	0.786*** (0.004)
Period=12	-0.071*** (0.003)	-0.077*** (0.009)		0.009 (0.005)	-0.064*** (0.010)	
Period=13	-0.123*** (0.003)	-0.154*** (0.009)		-0.084*** (0.005)	-0.144*** (0.010)	
Period=14	-0.127*** (0.003)	-0.180*** (0.008)		-0.136*** (0.005)	-0.178*** (0.009)	
Period=2 x *log_dist*	-0.851*** (0.006)	-1.283*** (0.021)	-1.364*** (0.025)	-1.436*** (0.008)	-1.547*** (0.025)	1.107*** (0.020)
Period=3 x *log_dist*	-0.776*** (0.004)	-1.075*** (0.014)	-1.396*** (0.015)	-1.326*** (0.006)	-1.301*** (0.017)	-1.133 (0.012)
Period=4 x *log_dist*	-0.298*** (0.003)	-0.881*** (0.013)	-1.136*** (0.011)	-0.572*** (0.004)	-1.069*** (0.015)	-0.924*** (0.009)
Period=5 x *log_dist*	-0.003 (0.002)	-0.222*** (0.008)	-0.943*** (0.004)	-0.062*** (0.003)	-0.271*** (0.009)	-0.781*** (0.003)
Period=6 x *log_dist*	0.015*** (0.002)	-0.133*** (0.007)	-0.871*** (0.004)	-0.022*** (0.003)	-0.164*** (0.008)	-0.721*** (0.003)
Period=7 x *log_dist*	-0.017*** (0.002)	-0.134*** (0.007)	-0.844*** (0.005)	-0.070*** (0.003)	-0.163*** (0.008)	-0.695*** (0.004)
Period=8 x *log_dist*	-0.006*** (0.002)	-0.078*** (0.006)	-0.934*** (0.004)	-0.043*** (0.003)	-0.095*** (0.007)	-0.772*** (0.004)
Period=9 x *log_dist*	-0.005** (0.002)	-0.039*** (0.006)	-0.955*** (0.004)	-0.041*** (0.003)	-0.048*** (0.007)	-0.789*** (0.003)
Period=10 x *log_dist*	-0.014*** (0.002)	-0.062*** (0.006)	-1.050*** (0.005)	-0.055*** (0.003)	-0.076*** (0.007)	-0.865*** (0.004)
Period=11 x *log_dist*	0.020*** (0.002)	0.022*** (0.005)	-0.999*** (0.004)	0.013*** (0.003)	0.025*** (0.006)	-0.828*** (0.003)
Period=12 x *log_dist*	-0.026*** (0.002)	-0.032*** (0.006)		-0.071*** (0.003)	-0.041*** (0.007)	

Period=13 x *log_dist*	0.002 (0.002)	-0.021*** (0.006)		-0.018*** (0.003)	-0.026*** (0.007)	
Period=14 x *log_dist*	0.040*** (0.002)	0.008 (0.005)		0.048*** (0.003)	0.008 (0.006)	
log_dist				-1.008*** (0.001)	-1.134*** (0.001)	-1.281*** (0.001)
Month FE	Included	Included	Included	Included	Included	Included
Day of the Week FE	Included	Included	Included	Included	Included	Included
District FE	Included	Included	Included	Not Included	Not Included	Not Included
Observations	823,439,472	190,177,776	1,003,696,765	823,439,472	190,177,776	1,003,696,765
Log-likelihood	-67,887,096	-14,181,463	-42,604,579	-76,159,339	-14,814,655	-43,098,465
Pseudo R2	0.163	0.094	0.079	0.061	0.054	0.068
AIC	135,774,343	28,363,045	85,209,378	152,318,767	29,629,400	86,197,009
BIC	135,775,752	28,364,069	85,211,438	152,319,601	29,630,168	86,197,739

Note. Model 1 and 2 are estimated using a new variable for period defined as; *Period*=1 ("Pre-covid"), *Period*=2 (last half of March 2020), *Period*=3 for the month of April 2020 and so on.
*$p<0.05$, **$p<0.01$, ***$p<0.001$.