CONTRIBUTI DI RICERCA CRENOS

A HUFF MODEL WITH HETEROGENEOUS RETAILERS FITS WELL IN SOUTHERN ITALY

Massimo Del Gatto
Carlo S. Mastinu

WORKING PAPERS

2017/11

CUEC

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENoS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.it
info@crenos.it

CRENoS – CAGLIARI
VIA SAN GIORGIO 12, I-09100 CAGLIARI, ITALIA
TEL. +39-070-6756406; FAX +39-070- 6756402

CRENoS - SASSARI
VIA MURONI 23, I-07100 SASSARI, ITALIA
TEL. +39-079-213511

Title: A HUFF MODEL WITH HETEROGENEOUS RETAILERS FITS WELL IN SOUTHERN ITALY

ISBN: 978 88 9386 050 5

First Edition: November 2017

A Huff model with heterogeneous retailers fits well in Southern Italy[1]

Massimo Del Gatto[2]
'G.d'Annunzio' University
and
CRENoS

Carlo S. Mastinu
University of Cagliari
and
CRENoS

Abstract

In a Huff model with heterogeneous retailers, a higher market potential within the trade area should result in higher average productivity and lower productivity dispersion, through the selection of the best stores. Using a unique dataset encompassing 14,212 Italian retailers, we report evidence of such a process at the municipality level. The effectiveness of this process is widespread in Southern Italy but not in Northern and Central Italy, suggesting the selection dynamics to be strongly affected by context factors related to an upper geographical scale. Interestingly, the extent of provincial/regional accessibility is not among these factors. This evidence is robust to controlling for local context factors such as financial risk and entry restrictions. Notably, entry restrictions are found to enhance selection.

Keywords: Huff model, firm selection, accessibility, trade areas, retail location
JEL Classification: R12, F12, R3, L81.

[1] The authors gratefully acknowledge financial support from the 'Analisi dei costi economici addizionali attribuibili allo stato di insularità, con particolare riferimento alla di differenza rispetto a casi di "geographic remoteness" riconosciuti nell'ambito della politica regionale europea', a project of European interest funded by the local government of the Sardinia region of Italy [grant number CRP-27162].

[2] Corresponding author: Massimo Del Gatto, `G.d'Annunzio' University, Department of Economics (DEc), Viale Pindaro 42, 65127 - Pescara (Italy). massimo.delgatto@unich.it

1 Introduction

The study of retail trade areas has a long tradition. According to Reilly's (1929, 1931) "law of retail gravitation", the market potential of two competing retail stores depends on their relative size, on the one hand, and their relative distance from potential customers, on the other hand.

Huff (1962, 1963) contributed to the quantification of retail trade areas by modeling the probability that consumers patronize different competing stores within the same area. According to the Huff model, this probability is a function of the stores' accessibility, relative to its competitors, and can be estimated through gravity regressions. The popularity and longevity of this approach can be attributed to its comprehensibility, relative ease of use, and its applicability to a wide range of problems. In recent years, several generalizations have been developed (e.g., Nakanishi and Cooper, 1974 and 1982), but the logic of the model remained basically unchanged. In particular, with the development and diffusion of new methodologies, based on the use of Geographic Information Systems (GIS), the model has been greatly extended and enriched (Birkin, 1995; Satani et al., 1998; Huff, 2003; Suárez-Vega et al., 2011; Xu and Liu, 2004) and is now the main tool used by retailers in choosing the location of their stores.

However, while the Huff model is regarded as a cornerstone approach in both geographical (e.g., Kwan, 1998; Karst and Van Eck, 2001) and marketing (e.g., Bell and Tang, 1998; Grewal et al., 2009) literature, it received relatively little attention from economic literature both at the theoretical and empirical level, with only a few notable exceptions such as Davis (2006).

From this point of view, a dimension which is completely neglected is the potential process of firm selection, associated with competition and market size, stemming from differences in productivity across firms/stores.

In fact, as advocated by the "New New Trade Theory" (hereafter, NNTT) models[3], the size of the market can be associated to selection effects stemming from higher factor market (e.g., Melitz, 2003) and/or product market (e.g., Melitz and Ottaviano, 2008; Corcos et al., 2012) competition. According to this type of selection, high-productivity firms succeed in charging

[3] The expression "New Trade Theory" was coined to refer to a strand of international trade literature, pioneered by Krugman (1980) and furtherly developed by Dixit and Norman (1980), Markusen (1981), Helpman (1984) and Helpman and Krugman (1985) among others, focusing on the role of increasing returns to scale and imperfect competition in international trade. While New Trade models successfully explained some key facts in international trade, such as the emergence of intra-industry flows, subsequent literature highlighted additional competition effects: higher competition forces the least productive firms to leave the market (Bernard and Jensen, 1999; Aw, Chung and Roberts, 2000; Clerides et al., 1998) and induces market share reallocations towards the more productive firms (Pavcnik, 2002, Bernard, Jensen and Schott 2003). Recent theoretical literature accommodated this "selection effect" by enriching the New Trade Theory approach with the assumption that firms are heterogeneous in terms of productivity (i.e. total factor productivity). This generated the class of models [i.e. Bernard et al. (2003), Melitz (2003), Ottaviano and Melitz (2004)], referred to as "New New Trade Theory".

lower prices for goods of a given quality, or in offering goods of superior quality at given prices. This allows them to gain market share at the expense of the less productive firms. Through this process, the better firms earn handsome profits, the mediocre ones lower profits, and the worst soon disappear, being unable to cover their production costs with revenues.[4] This reasoning finds ample empirical support for the manufacturing sector (see e.g., Pavcnik, 2002; Bustos, 2011), but its relevance for the retail sector has not been documented so far, at least to our knowledge.

To emphasize this aspect, we incorporate heterogeneous (in terms of productivity and, hence, marginal costs) retailers into the Huff (1963) model, to show that when only size and distance, as well as consumers' income, matter for demand (demand is inelastic), the process of firm selection is tougher when the market potential in the area is higher (that is, the larger is the total available income that can be reached and the smaller is total size of the competitors). Land is costly, and less productive firms can only afford a relatively smaller selling area. The presence of entry costs imposes a threshold, in terms of size. A larger local market, by increasing the profit maximizing size, lowers the level of costs above which firms are not able to serve the market, thereby decreasing both the average and the dispersion of the marginal costs distribution in the municipality. Accordingly, trade areas with a higher market potential should be characterized by higher average productivity (i.e., lower costs) and lower productivity dispersion (i.e., lower marginal cost dispersion).

We apply this model by taking advantage of a unique dataset, encompassing information on balance sheet items, size (in square meters) and geographical coordinates for 14212 Italian retailers (hypermarkets, supermarkets, discounters, and small retailers). We compute a theory-based relative measure of market potential (the ratio of distance-weighted consumers' income to distance-weighted store size in the trade area) at the municipality level which, according to the model, should be negatively correlated with the first and second moments of the cost distribution.

Indeed, we find evidence of such relationships in general. Moreover, when we take the geographical articulation (North, Central, South) into account, in order to control for a number of factors (e.g., infrastructure, institutional quality and regional autonomy, local regulation, financial institutions, labor market thickness, and human capital, among others) that are likely to affect the effectiveness of the selection process at the local level, we show this evidence to be pervasive in Southern Italy. This evidence is robust to controlling for local factors such as financial risk and entry barriers in the retail sector. Notably, the latter is found to foster the process of store selection.

While documenting that the selection effect also takes place at the local level (for non-tradables), our results point out that the effectiveness of the selection process is significantly affected by context factors related to an upper geographical scale. As this aspect has not been

[4] The productivity of firms that generate revenues barely sufficient to cover costs defines the threshold below which it is impossible for a firm to survive in the market. This threshold of survival determines the average productivity of active firms.

highlighted earlier for the retail sector, it might deserve further attention, as the identification of the factors affecting the pervasiveness of the selection process at the local level might provide the public authority with important policy implications. While we leave this issue for further research, we show that the extent of provincial or regional accessibility is not one of those context factors.

The article articulates as follows. In sections 2 and 3, we present the model and the empirical strategy. In section 4 we describe the data. Benchmark results are reported in section 5, followed by a number of robustness checks implemented in section 6. Section 7 concludes.

2 Model

In the original Huff model, the probability P_{if} that a consumer i, located in trade area A, will select store f located in site f (letter f is used to refer to both the store and its location), among all possible alternatives in A, is assumed to be a positive function of the store f's sales area (i.e., interchangeably referred to as size hereinafter) s_f and a negative function of its distance τ_{if} from the consumer (indexed i). Size and distance are evaluated in relative terms with respect to all possible alternative stores in A. To represent an alternative, a store has to fall within a given traveling time to the consumer. Hence, the probability can be written as

$$P_{if} = \frac{s_f^{\alpha} \tau_{if}^{\beta}}{\sum_{r=1}^{R(A)} s_r^{\alpha} \tau_{ir}^{\beta}}, \quad (1)$$

where index r identifies a general retailer (i.e., store), and $R(A)$ denotes the number of retailers in A.[5] Since demand is inelastic, the total demand available to store f can be expressed as

[5] The demand function in equation (1) is based on the assumption that the probability that consumer i, confronted with a set of alternatives, will select a given store is directly proportional to the perceived utility of each alternative. The choice is probabilistic, and each store is characterized by a positive probability $P_{if} = U_{if} / \sum_{r}^{R(A)} U_{ir}$ to be chosen, with U_{if} denoting the consumer's utility associated to choice f (and $\sum_r P_{ir} = 1$). Assuming that U_{if} is directly proportional to s_{if} and inversely proportional to τ_{if}, with the degree of proportionality governed by the two parameters α and β, yields the demand function in equation (1). Inelasticity entails that, if only one store existed, the total number of consumers would patronize it regardless of where it is located.

$$D_f = \sum_{i=1}^{I(A)} P_{if} B_i = s_f^{\alpha} \Phi_f$$
$$\text{with} \quad \Phi_f = \frac{\Psi_f}{\Theta_A}, \quad \Psi_f = \sum_{i=1}^{I(A)} B_i \tau_{if}^{\beta} \quad \text{and} \quad \Theta_A = \sum_{i=1}^{I(A)} \sum_{r=1}^{R(A)} s_r^{\alpha} \tau_{ir}^{\beta} \qquad (2)$$

where $I(A)$ is the number of consumers located in A and B_i is their income. As well as on its size, a store f's total demand depends on the two terms Ψ_f and Θ_A. The former is a distance-weighted measure of the total consumers' income accessible from location f, which varies across locations within the trade area. The latter is a distance-weighted measure of the total stores' size in the area and is therefore specific to the area. The ratio of these two terms (i.e., Φ_f) can be referred to as the *Available Market per Unit of Sales Area* (hereafter, AMUSA). This is a relative measure of (within-area) market potential which is comparable across trade areas.

This simple demand structure can be used to contextualize the original Huff model into a framework in which heterogeneous retailers choose their profit maximizing size. Heterogeneity is expressed in terms of inverse total factor productivity – i.e., Unit Input Requirement (UIR) – hereinafter referred to as c_f.

Retailers' activity only requires land (i.e., size), as a production factor. The production function is modeled as $Y_f = s_f^{\gamma}/c_f$. Firms use the same technology but differ in the UIR term. Markets (areas) are segmented, entailing that multi-store firms independently maximize profits from different locations, so that the decision concerning store f at location f can be always dealt with as the decision of a single-store firm with UIR equal to c_f.[6] Thus, the two terms 'store' and 'firm' can be used interchangeably. The marginal costs faced by firm f is ωc_f, with $\omega = (r)^{\gamma}$. Here, r denotes the rental price of land, which can be either specific or non-specific to the trade area (in the application we set it as province specific).

Firm-store f sets its size by solving

$$Max_{s_f} \quad D_f - \omega c_f s_f = s_f^{\alpha} \Phi_f - \omega c_f s_f \qquad (3)$$

taken as given the total sales area in A. First order conditions yield

[6] Alternatively, one can imagine production to also require labor, with the latter inelastically provided by consumers at unit wage (i.e., $Y_f = c_f^{-1} s_f^{\beta_s} l_f^{\beta_l}$).

$$s_f^* = \left[\frac{\omega c_f}{\alpha \Phi_f}\right]^{\frac{1}{\alpha-1}}. \qquad (4)$$

Under the standard assumption that $\alpha < 1$, a negative relationship between optimal store size and the firm's UIR emerges. In fact, the profit maximizing size is higher when the UIR is lower (the more productive is the firm), and the AMUSA is higher.

Free entry in the trade area imposes

$$\int_0^{\bar{c}_f} s_f^{\alpha} \Phi_f - \omega s_f c_f \; dG(c) = z_E \omega \qquad (5)$$

where $\bar{c}_f$ refers to the cutoff UIR level above which stores are not able to keep serving the market, and $z_E \omega$ is a fixed entry cost that each firm has to bear in order to open a new store. Also, z_E can be thought of as either specific or non-specific to the trade area. In the application, we assume it to be the same for all the Italian municipalities.

In order to solve equation 5 explicitly, let us assume, as in Melitz (2003) and Melitz and Ottaviano (2008), that the overall UIR distribution is Pareto, with shape parameter k and upper bound c_M within the support $[0, c_M]$: $G_M(c) = (c_f / c_M)^k$. By governing the probability to observe a store with UIR below a given level, C_M can be thought of as a parameter subsuming the exogenous differences in terms of the socio-economic context in which the areas are located. Examples can be the quality of infrastructure and financial institutions, or the local regulation concerning the retail sector. In the application we imagine the support of the UIR distribution to vary across the Italian macro-regions (North, Central, and South-Islands).

Using (4) to substitute for s_f, the solution of (5) provides us with the following expression for the UIR cut-off above which store f is not able to survive in area A:

$$\bar{c}_f = \Lambda \, z_E^{\frac{\alpha-1}{\upsilon}} \, \omega^{-\frac{1}{\upsilon}} \, \Phi_f^{\frac{1}{\upsilon}} \, c_M^{\frac{k(\alpha-1)}{\upsilon}}. \qquad (6)$$

where Λ is equal to $\dfrac{k+1}{k[\alpha^{-\alpha/(\alpha-1)} - \alpha^{1/(\alpha-1)}]}$ and $\upsilon = k(\alpha - 1) + \alpha$.[7]

Under the stability condition that $k > \dfrac{\alpha}{1-\alpha}$ (which implies that $\upsilon < 0$), the UIR threshold increases with the fixed cost of entry and with the land rent, and decreases with the

[7] Note that, given the one-to-one relationship between UIR and size, the equilibrium UIR no longer depends on size, once equation (4) is used to substitute for s_f into (5).

AMUSA.[8]

The rationale for the AMUSA effect is as follows. From equation (4), a higher AMUSA entails, for all firms, a higher profit maximizing size. However, land (i.e., size) is costly and the profit maximizing size is lower for retailers with a relatively high UIR. By imposing a minimum size, the entry cost also imposes a maximum possible UIR (i.e., UIR cut-off). Retailers with UIR values above this threshold level cannot afford a sufficiently large size (i.e. they have a too low profit maximizing size) and are not able to survive.

3 Testable implications and empirical strategy

Equation (6) yields testable implications that can be studied through the various moments of the UIR distribution characterizing the trade areas. In fact, with the UIR Pareto distributed, a lower UIR cut-off maps into lower measures of central tendency (e.g., mean and median) and lower measures of dispersion (e.g., standard deviation and IQ range).[9] In particular, the UIR average and standard deviation in trade area A are given by

$$\text{AVG}(c_f) = \frac{k}{k+1}\bar{c}_f \qquad \text{and} \qquad \text{SD}(c_f) = \frac{1}{\sqrt{k(k+2)}(k+2)}\bar{c}_f \qquad (7)$$

These two expressions reveal the aggregate effect of the selection process featured by the model: a higher AMUSA is associated, through (6) and (7), with lower UIR average and standard deviation in the area.

In the following analysis, we bring these predictions to data at the national level. For this purpose, we first use information at the single store level to estimate each store's UIR, and then aggregate at the municipality level. This amounts to setting f equal to the municipality. Hereafter, we use index m to refer to the municipality.

Under this strategy, our estimating equation is

$$ln(Q_m) = \lambda_0 + \lambda_1 ln(\Phi_m) + \lambda_2 ln(\omega_p) + \lambda_3 macroregion + \lambda_4 ln(\Phi_m) * macroregion + \varepsilon_m \qquad (8)$$

Depending on the specification, the dependent variable Q_m is the average or the

[8] The condition $k > \frac{\alpha}{1-\alpha}$ allows the integral in (5) to converge to a meaningful solution.

[9] The simplest measure of dispersion would be the 'range', defined as the UIR gap between the best- and worst-performing stores. However, given the support $[0, \bar{c}_f]$, the range is simply equal to $\bar{c}_f$, so it increases with the degree of competition in the trade area. While easily understood, being based on the two boundary values only, the range is necessarily very sensitive to extreme observations and should be used together with other measures. The 'standard deviation' is the most widely used measure of dispersion. Although less sensitive, the SD might also be problematic in highly skewed distributions. In section 6, we provide robustness checks for both average and dispersion, by relying on the median and the interquartile range.

standard deviation of the UIR distribution in municipality m (in section 6 we also present results for the median and the interquartile range). λ_0 captures the effect of the entry cost term $z_E^{\frac{\alpha-1}{\upsilon}}$ and the bundle of parameters Λ, which are assumed to be the same throughout Italy . The rental price of land ω is assumed (see Section 4) to be province-specific, which seems to be realistic. ε is an iid error term.

It is well-known that the Italian territory is characterized by huge differences in socio-economic features (infrastructure, institutional quality and regional autonomy, local regulation, financial institutions, labor market thickness, and human capital, among others). By interacting with local competition and market size (our variables of interest), these factors are likely to impact the effectiveness of the selection result. The model suggests a convenient way to deal with these factors. In fact, we can imagine them to condition the UIR distribution by affecting its upper bound C_M. As noticed, this parameter governs the probability to observe a store with a UIR below a given level. Since most of the heterogeneity in these characteristics is correlated with latitude, their action can be easily taken into account by including a vector of dummies (i.e ., $macroregion$) controlling for municipality belonging to one of the three Italian macro-regions: North, Central and South (with the two island regions, Sardinia and Sicily, included in the South). Under the assumption that C_M varies across the macroregions, these dummies should capture the exogenous differences in the UIR support. In order to isolate the differential effect of the AMUSA in the different macro-regions, an interaction term $AMUSA * macroregion$ is also included.

4 Data and variables definition

Our main explanatory variable is the AMUSA. To compute this term, we use a GIS software to calculate Ψ_m and Θ_A by aggregating over consumers, for Ψ_f, and stores, for Θ_A, within a traveling time of fifteen minutes. For each municipality, the trade area A consists of all the municipalities located within such traveling time.

Data on retailers are provided by Nielsen[10], which conducts a regularly updated census on Italian mass retailers. Our data, updated in September 2016, include 27,966 stores divided into four categories: hypermarkets, supermarkets, discounters, and small retailers. For each store, in addition to geographical coordinates, a number of variables are provided, including the size of the sales area, expressed in square meters. The geographic distribution of the stores is visualized in Figure 1.

The UIR index of each store is computed as the reciprocal of the difference (i.e., Solow residual) between the store's actual and predicted value added. To compute the latter, we estimate a standard Cobb-Douglas production function including labor and capital. The

[10] See: http://www.nielsen.com.

estimated coefficients amount to 0.301 (with standard error 0.0104), for capital, and 0.699 (with standard error 0.0189), for labor. To this aim, we match the Nielsen data with balance sheet data drawn from the AIDA database (provided by Bureau van Dijk). 14,212 (of the 27,965) stores were matched successfully.[11] In order to take simultaneity issues into account, UIR is estimated following the Olley-Pakes (1996) procedure. A detailed description of the estimation is reported in Appendix A.

For the term τ, an Origin–Destination (OD) matrix among Italian municipalities is needed. This is calculated using the entire network of the Italian extra-urban roads, updated to 2016. Our estimate of driving times are accomplished through the use of a GIS program and by taking into account four key variables: length, direction of travel, hierarchy of the functional road classes, and journey speed. In order to determine the journey speed in each class of road, we referred to the Ministerial Decree of November 5, 2001 (so-called Decreto Lunardi), which identifies 14 types of roads and assigns to each type a lower and an upper speed limit. We have taken the latter as the reference speed of travel.[12] Our OD matrix includes 8085 Italian municipalities and consists of the driving times that separate each municipality from the municipalities located within a travel time of up to 15 minutes. An alternative travel time area of 20 minutes has been used, finding no notable differences in the econometric results.

With the OD matrix in our hands, we calculate Ψ_f following the potential accessibility formulation proposed by Wegener et al. (2002): $\Psi_m = \sum_j^{M(A)} B_j\ exp(-\rho\tau_{jm})$. Where index m refers to the given municipality, index j refers to the generic municipality located in the area (with M(A) denoting the number of municipalities located within the driving distance of fifteen minutes from m), and τ_{jm} is the vector of journey time between municipality m and municipality j. Further, B_j is the total available income in municipality j, drawn from ISTAT (the Italian National Institute of Statistics). Also, ρ is a decay parameter set at 0.05. The set of the τ_{jm} for all the municipality pairs forms the OD matrix. The computed values of Ψ_m are reported in Figure 2.

A similar logic is followed to compute Θ_A, with municipalities' income replaced by the total sales area (expressed in square meters) available there, provided by Nielsen.

To measure the rental price of capital (r_A), we rely on data from the real estate market and consider the average price of sales and rents, downloaded (in September 2016) from a

---

[11] In the case of multi-store firms, the UIR refers to the main branch.

[12] Notice that the resulting OD matrix underestimates actual travel times, for different reasons. First, our data only includes the extra-urban roads, so we do not consider the time required to reach the extra-urban road network. Second, the analysis excludes any kind of barriers (such as traffic lights and toll gates). Third, we use the maximum allowed speed as reference speed of travel.

popular Italian real estate website (www.immobiliare.it), at the Nuts 3 level.

In the robustness analysis we also make use of a measure of entry barriers in the retail sector, calculated at the Nuts 3 level (i.e. Italian provinces) by Schivardi and Viviano (2011). The index is computed as the ratio of population to admissible floor space (PAFS), as regulated by law.[13] The higher this ratio, the greater the entry restrictions.

Finally, we also use an index of financial risk, compute at the Nuts 3 level (i.e. Italian provinces) by ISTAT (the Italian National Institute of Statistics) as the ratio of non-performing to performing loans granted to all types of firms.

Descriptive statistics for the main variables used in the analysis are reported in Table 1.

5 Results

Table 2 reports the results of the benchmark estimation of equation (8) for the central tendency effect - i.e., $Q_m = \text{AVG}(c_m)$ - and the dispersion effect - i.e., $Q_m = \text{SD}(c_m)$, in columns 1 to 3 and 4 to 6, respectively.

Our main variable of interest is the AMUSA. As an overall effect, the first row suggests a negative relationship with both the UIR average (column 1) and standard deviation (column 4). This is in line with the selection effect predicted by the model.

In columns 2 and 5 , we control for whether the municipality belongs to Northern, Central, or Southern Italy using the $North$ and $South$ dummies and setting the Central as the benchmark, so that the coefficients of the two dummies represent the differential AVG and SD effects in Northern and Southern Italy. In line with common wisdom concerning the productivity gap of Southern Italy, the UIR distributions in the South are characterized by higher average and dispersion.

To gain further insights on this dimension, in columns 3 and 6, the AMUSA is interacted with the $macroregion$ dummy. As Central Italy is the benchmark case, the coefficient in the first row represents the average AMUSA effect in the Central municipalities, while the coefficients of $AMUSA*North$ and $AMUSA*South$ pick the differential AMUSA effect in the Northern and Southern municipalities (with respect to the Central municipalities). Interestingly, we find that while the productivity gap is confirmed, most of the overall effect estimated in columns 1-2 and 4-5 has to be attributed to the Southern municipalities (i.e., the coefficient of $AMUSA*South$ is negative and significant, while the coefficients of $AMUSA$ and $AMUSA*North$, are both not significant).

Thus, the evidence in favor of a selection effect fostered by the local market potential is

[13] As noticed by Schivardi and Viviano (2011): the Italian retail sector, which has a prevalence of traditional small stores, underwent a major regulatory change in 1998. A central feature of the new law is that it delegates the regulation of entry of medium-large stores to local authorities. As it turns out, local regulations differ substantially in their approach to competition: in particular, most regions have established stringent ceilings to the floor space that can be authorized for entry of medium-large stores at the local level.

pervasive in the Southern municipalities and actually absent in the rest of Italy. This difference points to the existence of macro-regional characteristics affecting the pervasiveness of the selection effect at the local level. This is something that has not been highlighted before.

The differentials in in productivity dynamics (both firm-level and aggregate) across the Italian regions are well known and related to a number of factors. While examining the details of these factors is definitively beyond the scope of the present analysis (see however Calligaris et al., 2016 for a recent analysis), a dimension that is worth controlling for in our analysis is whether the different documented effects are explained by additional competition effects taking place at a 'less local' spatial scale: the province and/or the region.

To investigate this, we recompute the AMUSA at the level of the 103 Italian provinces using, for the numerator Ψ_f, the measure of multi-modal accessibility provided by the European Spatial Planning and Observation Network (ESPON) for the Italian provinces. The recomputed AMUSA (Φ_p) is included, together with the AMUSA computed at the municipality level (correlation is 0.0330), in the regressions reported in the upper panel of Table 3.[14] Although significant, the provincial AMUSA effect does not absorb the significance of the differential effect in Southern Italy (the interaction term $AMUSA * South$, still considered at the municipality level, remains significant). Further, also the overall selection effect documented in columns 1–2 and 4–5 of row 1 remains significant. To check for the same effect at the regional level, we use, in the bottom panel, the ESPON measure of multi-modal regional accessibility[15]. Neither the overall selection nor the differential effect characterizing the municipalities located in Southern Italy can be explained by the regional scale of accessibility.

6 Robustness

Robustness of the results in section 5 is checked in several ways.

As a first experiment, we ask whether the selection process characterizing Southern Italy can be explained by local context factors. To this aim, we perform again the regressions in Table 2 controlling for two province-level characteristics. First, an index of entry restrictions computed

[14] The ESPON accessibility measures used in Table 3 are province-level (upper panel) and region-level (bottom panel) variables computed on 2006 data. The accessibility of province/region j is defined as $Acc_j = \sum_r Z_r exp(-\beta \bar{c}_{jr})$, where $\bar{c}_{jr}$ refers to the aggregation, over transport modes (i.e., air, rail, road), of the cost (c_{jrt}) of reaching r from j using transportation mode t - i.e., $\bar{c}_{jr} = -(1/\lambda) ln \sum_t exp(-\lambda c_{jrt})$, where Z_r is GDP-PPS per capita and population in region r, respectively, for the two measures computed at the province and region level, and λ is a parameter indicating the sensitivity to travel cost. The interpretation is that the accessibility of j increases with the number of "accessible" provinces/regions and with their size (either GDP or population).

[15] In this case, we do not divide by Θ_A.

as the ratio of population to admissible floor space (PAFS), as regulated by the Italian law. The higher this ratio, the greater the entry restrictions. Second, we control for the degree of financial risk measured through the incidence of non-performing loans (i.e. the ratio of non-performing to performing loans). Arguably, the higher this index, the more costly is capital at the level of the local financial system. The output of these regressions is reported in Table 4. Entry restrictions significatively contribute to the selection process by lowering both the first and the second moments of the marginal cost distribution. Financial risk displays a mild positive correlation with productivity dispersion. The evidence on the selection effect associated to AMUSA is unaffected.

As a second check, we use alternative measures of central tendency and dispersion measures. In fact, the mean is not a good measure of central tendency in skewed distributions (like the Pareto for $k > 1$). For such distributions, the median is a better measure, and in our model it is equal to $\mathrm{MDN}(c_f) = (0.5)^{\frac{1}{k}} \bar{c}_f$. For dispersion, a better measure in skewed distributions, compared to the standard deviation, can be the “interquartile range", defined as the difference between the 75th and the 25th percentiles. In our model, this is equal to $\mathrm{IQ}(c_f) = \left[(0.75)^{\frac{1}{k}} - (0.25)^{\frac{1}{k}} \right] \bar{c}_f$. Both these measures are increasing in the UIR cutoff. The results of the alternative regressions including the median and the IQ range are reported in columns 1-3 and 4-6 of Table 5, respectively. The benchmark results are confirmed .

In our third robustness check, we recognize that municipalities with highest accessibility are Rome and Milan, located in Central and Northern Italy, respectively. If the selection effect is particularly low in these two provinces, we are likely to estimate a not significant AMUSA effect for the Central and Northern municipalities. The regressions in Table 7, run without the municipalities located in the provinces of Rome and Milan, show that this is not the case.

Our dataset covers four categories of stores (hypermarkets, supermarkets, discounters, and small retailers). In principle, the results might differ across categories. In particular, one might think of the market potential of supermarkets, discounters, and small retailers as associated to a smaller geographical scale. In Table 8 we show that the benchmark results remain valid when only hypermarkets are considered.

Finally, the AMUSA is a composition of distance-weighted consumers’ income and distance-weighted store size (i.e., Ψ_f and Θ_A respectively). In principle, one would expect these two terms to be in a negative and positive relationship with the UIR average and dispersion, respectively. In Table 6, we consider the two variables separately and show that this is indeed the case.

7 Conclusions

According to NNTT models, a higher market size, through its induced competition effects, allows high-productivity firms to earn substantial profits at the expense of the less productive firms, which are forced to leave the market being unable to cover their production costs with revenues. While this reasoning finds ample empirical support for tradable goods (i.e., the manufacturing sector), its relevance for non-tradables (e.g., the retail sector) has not been documented so far.

To emphasize this aspect, we incorporated heterogeneous retailers into the original Huff (1963) model, the cornerstone model in retail location analysis. Since demand is inelastic, only size and distance, as well as consumers' income, matter for demand. On the one hand, land is costly and less productive firms can only afford a relatively smaller selling area; on the other hand, entry costs impose a threshold, in terms of size. A larger local market, by increasing the profit maximizing size, lowers the level of costs above which firms are unable to keep serving the market, thereby decreasing both the average and the dispersion of the marginal costs distribution in the municipality.

We applied this concept to the data taking advantage of a unique dataset encompassing 14,212 Italian retailers (hypermarkets, supermarkets, discounters, and small retailers). We computed a theory-based relative measure of market potential (the ratio of distance-weighted consumers' income to distance-weighted store size in the trade area) at the municipality level. According to the model, this measure should be negatively correlated with the first and second moments of the cost distribution, through the selection of the best stores.

As well as providing overall evidence of such a relationship in Italy, we took a geographical perspective (i.e., belonging to Northern, Central or Southern Italy) to control for factors (e.g., infrastructure, institutional quality and regional autonomy, local regulation, financial institutions, labor market thickness and human capital, among others) that are likely to condition the effectiveness of the selection process at the local level. We found that the evidence concerning the local selection effect is pervasive in Southern Italy and absent in the Northern and Central municipalities. This evidence is robust to controlling for local factors such as financial risk and entry barriers in the retail sector. Notably, higher entry restrictions are associated with tougher selection.

These findings are new in retail literature and suggest that: i) a higher market potential at the retail trade area level can be associated to aggregate productivity advantages, fostered by a process of firm selection; ii) the selection process at the local level is significantly affected by context factors, external to the firms, which are related to an upper geographical scale. While the identification of these factors might provide public authorities with key policy messages, a suggestion we leave for future research, we show that neither regional nor provincial accessibility is among the latter.

References

Aw, B., Chung, S. and Roberts, M. (2000). 'Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China)', *World Bank Economic Review*, 14, 65-90.

Bell, D. R., Ho, T. H., and Tang, C. S. (1998). Determining where to shop: Fixed and variable costs of shopping. *Journal of Marketing Research*, 352-369.

Bernard, A. and Jensen, B. (1999). 'Exceptional Exporter Performance: Cause, Effect, or Both?', *Journal of International Economics*, Vo. 47, 1-25.

Bernard, A., Jensen, B. and Schott, P. (2006). 'Trade costs, firms and productivity', *Journal of Monetary Economics*, 53, 917-37.

Bernard, A., Eaton, J., Jensen, J. and Kortum, S. (2003). 'Plants and Productivity in International Trade', *American Economic Review*, 93, 1268-1290.

Birkin, M. (1995). *Customer targeting, geodemographics and lifestyle approaches*. In P. Longley, and G. Clark (Eds.), GIS for business and service planning. Cambridge, U.K: Geoinformation International.

Bustos, P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms. *American Economic Review*, 101, 304-340.

Cooper, Lee G., and Masao Nakanishi (1988). *Market-Share Analysis* (Massachusetts: Kluwer Academic Publishers).

Corcos, G., Del Gatto, M., Mion, G. and Ottaviano, G.I.P. (2012). 'Productivity and firm-selection: quantifying the new gains from trade', *The Economic Journal*, 122(561), 754-798.

Clerides, S., Lach, S. and Tybout, J. (1998). 'Is Learning by Exporting Important? Micro-dynamic Evidence from Colombia, Mexico, and Morocco', *Quarterly Journal of Economics*, 113, 903-947.

Davis, P. (2006). Spatial competition in retail markets: movie theaters. *The RAND Journal of Economics*, 37(4), 964-982.

Dixit, A., and Norman, V. (1980). *Theory of international trade: A dual, general equilibrium approach*, Cambridge University Press.

Geurs, K. T., and J. R. Ritsema van Eck (2001). Accessibility measures: review and applications. Evaluation of accessibility impacts of land-use transportation scenarios, and related social and economic impact, mimeo.

Grewal, D., Levy, M., and Kumar, V. (2009). Customer experience management in retailing: An organizing framework. *Journal of Retailing*, 85(1), 1-14.

Helpman, E. (1984). Increasing returns, imperfect markets, and trade theory. *Handbook of international economics*, 1, 325-365.

Helpman, E., and Krugman, P. R. (1985). *Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy*, MIT press.

Huff, D. L. (1962). A note on the limitation of interurban gravity models. *Land Economics*, 38(1), 64-66.

Huff, D. L. (1963). Probabilistic Analysis of Shopping Center Trade Areas. *Land Economics*, 39, 81-90.

Huff, D. L. (2003). *Parameter estimation in Huff model*, www.esri.com/news/arcuser/1003/files/huff.pdf.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5), 950-959.

Kwan, M. P., and Weber, J. (2008). Scale and accessibility: Implications for the analysis of land use-travel interaction. *Applied Geography*, 28(2), 110-123.

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

Markusen, J. R. (1981). Trade and the gains from trade with imperfect competition. *Journal of international economics*, 11(4), 531-551.

Melitz, M. (2003). 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica*, 71, 1695-1725.

Melitz, M. and Ottaviano, G.I.P. (2008). 'Market size, trade, and productivity', *Review of Economic Studies*, 75, 295-316.

Nakanishi, M., and Cooper, L. G. (1974). Parameter estimation for a multiplicative competitive interaction model: least squares approach. *Journal of Marketing Research*, 303-311.

Nakanishi, M., and Cooper, L. G. (1982). Technical Note-Simplified Estimation Procedures for MCI Models. *Marketing Science*, 1(3), 314-322.

Nakanishi, Masao, and Lee G. Cooper (1982). Simplified Estimation Procedures for MCI Models, *Marketing Science*.

Pavcnik, N. (2002). Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants, *Review of Economic Studies*, 69, 245-276.

Reilly, W. J. (1929). Methods for the study of retail relationships. University of Texas, Bulletin No 2944.

Reilly, W. J. (1931). *The law of retail gravitation*. New York: Knicker-bocker Press.

Satani, N., Uchida, A., Deguchi, A., Ohgai, A., Sato, S., and Hagishima, S. (1998). Commercial facility location model using multiple regressions analysis. *Computer, Environment, and Urban Systems*, 22(3), 219-240.

Schivardi, F., and Viviano, E. (2011). Entry barriers in retail trade. *The Economic Journal*, 121(551), 145-170.

Suàrez-Vega, R., Santos-Peñate, D., Dorta-Gonzàlez, P., and Rodrìguez-Dìaz, M. (2011). A multi-criteria GIS based procedure to solve a network competitive location problem. *Applied Geography*, 31, 282e291.

Wegener, M., Eskelinnen, H., Fürst, F., Schürmann, C., Spiekermann, K. (2002). *Criteria for the Spatial Differentiation of the EU Territory: Geographical Position*. Forschungen 102.2, Bonn: Bundesamt für Bauwesen und Raumordnung.

Appendix A: UIR estimation

Following the common approach to total factor productivity (i.e., tfp) estimation, the inverse of our UIR measure (i.e., $1/c_{ft}$) denotes store f 's tfp, so that our estimating production can be written (in logs) as

$$y_{ft} = \frac{1}{c_{ft}} + \mathbf{x}_{ft}\beta + e_{ft}. \tag{9}$$

$\mathbf{x}_{ft}$ is a $(1 \times N)$ vector of inputs, and β is the $(N \times 1)$ vector of the elasticities of output with respect to each input. The stochastic disturbance e_{ft} is meant to capture measurement errors and unobserved idiosyncratic shocks, due for instance to unanticipated (and thus uncorrelated with $\mathbf{x}_{ft}$) market changes.

The value of c_{ft} can be recovered by: i) estimating the vector $\hat{\beta}$; ii) computing the fitted value of firm f 's output $\hat{y}_{ft} = \mathbf{x}_{ft}\hat{\beta}$; iii) deriving $\hat{c}_{ft}$ as the inverse of (the exponential of) the difference between y_{ft} and $\hat{y}_{ft}$ – that is $\hat{c}_{ft} = \left[exp\left(y_{ft} - \mathbf{x}_{ft}\hat{\beta}\right)\right]^{-1}$.

In the application, we consider a two input case with capital measured through the book value of tangible fixed assets and labor measured through employment (data drawn from the AIDA – Bureau van Dijk database). Our production function $Y_f = s_f^{\gamma}/c_f$ is nested in this standard specification as far as stores' sales area is included in the book value of capital.

In principle the production function estimation could be carried out through a simple OLS regression. However, information on c_{it}, although unknown to the econometrician, is commonly used by the firm in its decision concerning the amount of inputs. This *simultaneity* issue makes the error term e_{ft} correlated with $\mathbf{x}_{ft}$ and the OLS–estimated $\hat{\beta}$ biased. In econometric parlance, c_{ft} is said to " transmit" to the explanatory variables, hence the expression " transmission bias".[16]

[16] Note that such bias cannot be removed by assuming that the productivity component is not observed by the firm, since in any case one has to reckon with the fact that the amount of inputs is jointly determined with y_{ft}, which is just an alternative way of saying that the error term is correlated with the explanatory variables. Whether we want to look at this correlation from the former or the latter point of view, y_{ft} and $\mathbf{x}_{ft}$ must be regarded as the solution of a simultaneous-equations system. Thus, the problem is one of *simultaneity*. Although these two ways of looking at simultaneity are equivalent with

The theoretical stratagems to which one can resort, in order to keep into account the presence of simultaneity, go along with the " anatomy" of the TFP component. Specifically, note that the unobserved (by the econometrician) tfp term in eq. (9) is both store-specific (the f index), and time-varying (the t index).

Traditional cross-section analysis (Douglas, 1948) substitutes a constant for the unobserved component (that is $c_{it} \rightarrow c$), so that all its variability is included in the error term and all the simultaneity bias passed on the estimates.

If panel data are available (i.e. data on the various stores for more than one year), a first step towards mitigating the simultaneity bias can be made by reducing c_{it} to a firm-specific (but time-invariant) unobserved effect ($c_{it} \rightarrow c_i$). In this case, the UIR component is understood as an unobservable effect in a *fixed-effects* estimation. While this approach takes account of the individual heterogeneity between firms, it still neglects the temporal dimension.

The semi-parametric methodology suggested by Olley and Pakes (1996) allows to keep also the time dimension into consideration. The approach consists of identifying a (proxy) variable that reacts to the changes in the UIR (or tfp) observed by the firm and is therefore a function of it. Insofar as this function proves to be invertible, its inverse can be calculated and plugged in the estimating equation before proceeding to estimate the production function parameters. Summing up, the idea behind this *proxy-variable method* consists of recovering the productivity component by the traces it leaves in the observed behaviour of the firm. This approach, firstly proposed by Olley and Pakes (1996) using investment as a proxy, has been extended by Levinsohn and Petrin (2003) to the use of the intermediate inputs (see also Ackerberg et al., 2006; Wooldridge, 2009; Gandhi et al., 2013, as well as Del Gatto et al., 2011 and Van Beveren, 2012 for surveys). Our implementation strictly follows the description of the Olley and Pakes procedure reported in Del Gatto et al. (2011), section 5.2.1.

Additional bibliography

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6), 2411-2451.

Del Gatto, M., Di Liberto, A., and Petraglia, C. (2011). Measuring Productivity. *Journal of Economic Surveys*, 25(5), 952-1008.

---

respect to the econometric stratagems to which one can resort, it is worth noting how the former poses, more properly, a problem of " omitted variables".

Gandhi, A., Navarro, S., and Rivers, D. A. (2011). On the identification of production functions: How heterogeneous is productivity? *mimeo*

Levinsohn, J. and Petrin, A. (2003) Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies* 70(2): 317-341.

Van Beveren, I. (2012). Total Factor Productivity Estimation: A Practical Review. *Journal of Economic Surveys*, **26**(1), 98-128.

Wooldridge, J. (2009). On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables. *Economics Letters*, **104**(3), 112-114.

Table 1: Descriptive statistics: main variables by region.

Region	$AVG(c_f)$	$Min(c_f)$	$Max(c_f)$	$SD(c_f)$	AMUSA (Φ_m)	Ψ_m	Θ_A	ω_p
ABRUZZO	0.0374	0.0007	0.0999	0.0253	58.77	204.83	6.50	6.14
BASILICATA	0.0611	0.0137	0.6441	0.0705	44.68	76.84	2.48	4.87
CALABRIA	0.1412	-0.2015	1.8552	0.1449	33.99	100.26	4.80	4.49
CAMPANIA	0.0544	-1.6419	0.6368	0.0465	44.54	908.76	24.22	5.69
EMILIA-ROMAGNA	0.0491	0.0007	0.1122	0.0297	59.88	936.80	17.05	7.99
FRIULI-VENEZIA GIULIA	0.0542	0.0138	0.1355	0.0276	36.46	686.92	21.97	6.83
LAZIO	0.0241	-6.3295	0.8734	0.0460	49.58	470.54	11.14	8.38
LIGURIA	0.0451	0.0007	0.1023	0.0302	59.99	421.04	8.15	9.25
LOMBARDIA	0.0553	-0.5898	0.8792	0.0365	53.72	2262.51	42.94	8.22
MARCHE	0.0572	0.0008	0.1168	0.0349	40.37	342.19	10.66	7.07
MOLISE	0.0434	0.0010	0.0744	0.0247	83.94	89.58	2.25	4.96
PIEMONTE	0.0139	-12.9399	1.7206	0.0939	66.41	689.02	16.32	6.26
PUGLIA	0.0178	-5.6128	0.3959	0.0517	43.24	597.55	19.38	6.06
SARDEGNA	0.0892	-0.2805	1.4740	0.0807	30.79	156.90	6.61	8.12
SICILIA	0.0362	-7.9466	1.2161	0.1019	34.75	387.54	12.61	5.73
TOSCANA	0.0752	0.0007	0.1636	0.0338	63.41	713.53	11.88	10.12
TRENTINO-ALTO ADIGE	0.0790	0.0007	0.1225	0.0224	54.89	225.09	4.98	9.87
UMBRIA	0.0582	0.0242	0.1168	0.0330	36.57	316.03	9.77	6.21
VALLE D'AOSTA	0.0376	0.0009	0.0864	0.0033	61.62	148.23	3.29	8.13
VENETO	0.0467	-0.3013	0.1646	0.0258	42.21	1057.92	29.20	7.99
Total	0.0540	-3.0552	0.7973	0.0491	50.81	875.28	19.63	7.24

Table 2: Benchmark results.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	$AVG(c_f)$	$AVG(c_f)$	$AVG(c_f)$	$SD(c_f)$	$SD(c_f)$	$SD(c_f)$
AMUSA (Φ_m)	-0.105***	-0.061*	0.062	-0.245***	-0.186***	-0.014
	(0.03)	(0.04)	(0.07)	(0.07)	(0.07)	(0.11)
Rental price of land (ω_p)	0.034	0.181***	0.152**	-0.204*	-0.081	-0.130
	(0.06)	(0.06)	(0.06)	(0.11)	(0.11)	(0.11)
North (dummy)		-0.121***	0.106		-0.249***	-0.218
		(0.04)	(0.33)		(0.06)	(0.56)
South (dummy)		0.101***	0.938***		0.013	1.438***
		(0.04)	(0.29)		(0.07)	(0.53)
AMUSA*North			-0.062			-0.010
			(0.09)			(0.15)
AMUSA*South			-0.239***			-0.416***
			(0.08)			(0.15)
Constant	-2.705***	-3.129***	-3.527***	-2.375***	-2.722***	-3.246***
	(0.14)	(0.16)	(0.24)	(0.26)	(0.30)	(0.37)
N	3656	3656	3656	1982	1982	1982
adj R^2	0.002	0.013	0.014	0.015	0.026	0.029

Standard errors are in parentheses. All variables are in logs.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Robustness: regressions with AMUSA at the provincial and regional level.

	(1) $AVG(c_f)$	(2) $AVG(c_f)$	(3) $AVG(c_f)$	(4) $SD(c_f)$	(5) $SD(c_f)$	(6) $SD(c_f)$
AMUSA (Φ_m)	-0.106***	-0.063*	0.086	-0.223***	-0.174**	-0.046
	(0.03)	(0.04)	(0.07)	(0.07)	(0.07)	(0.11)
AMUSA-province (Φ_p)	-0.007	-0.041**	-0.045***	0.083***	0.056**	0.050*
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Rental price of land (ω_p)	0.031	0.171***	0.137**	-0.190*	-0.063	-0.109
	(0.06)	(0.06)	(0.06)	(0.11)	(0.11)	(0.11)
North (dummy)		-0.179***	0.162		-0.169**	-0.325
		(0.05)	(0.33)		(0.07)	(0.56)
South (dummy)		0.065	1.008***		0.071	1.295**
		(0.04)	(0.30)		(0.07)	(0.54)
AMUSA*North			-0.094			0.039
			(0.09)			(0.15)
AMUSA*South			-0.269***			-0.361**
			(0.09)			(0.16)
Constant	-2.717***	-3.169***	-3.659***	-2.261***	-2.708***	-3.092***
	(0.14)	(0.16)	(0.25)	(0.27)	(0.30)	(0.38)
N	3656	3656	3656	1982	1982	1982
adj R^2	0.002	0.014	0.017	0.021	0.027	0.031

AMUSA (Φ_m)	-0.100***	-0.063*	0.061	-0.205***	-0.168**	-0.007
	(0.04)	(0.04)	(0.07)	(0.07)	(0.07)	(0.11)
Regional Accessibility	0.020	-0.013	-0.014	0.096***	0.061**	0.056*
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Rental price of land (ω_p)	0.041	0.180***	0.151**	-0.175	-0.066	-0.114
	(0.06)	(0.06)	(0.06)	(0.11)	(0.11)	(0.11)
North (dummy)		-0.141***	0.078		-0.151*	-0.134
		(0.05)	(0.32)		(0.08)	(0.57)
South (dummy)		0.092**	0.933***		0.067	1.420***
		(0.04)	(0.29)		(0.07)	(0.53)
AMUSA*North			-0.060			-0.008
			(0.09)			(0.15)
AMUSA*South			-0.241***			-0.397***
			(0.08)			(0.15)
Constant	-2.649***	-3.167***	-3.568***	-2.155***	-2.612***	-3.120***
	(0.14)	(0.16)	(0.24)	(0.27)	(0.31)	(0.38)
N	3656	3656	3656	1982	1982	1982
adj R^2	0.003	0.012	0.014	0.023	0.028	0.031

Standard errors are in parentheses. All variables are in logs.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Robustness: regressions controlling for admissible floor space (PAFS) and financial risk.

	(1)	(2)	(3)	(4)	(5)	(6)
	$AVG(c_f)$	$AVG(c_f)$	$AVG(c_f)$	$SD(c_f)$	$SD(c_f)$	$SD(c_f)$
AMUSA (Φ_m)	-0.048	-0.009	0.112*	-0.261***	-0.223***	0.015
	(0.03)	(0.04)	(0.07)	(0.07)	(0.07)	(0.12)
Rental price of land (ω_p)	0.087	0.199***	0.159**	0.005	0.085	0.030
	(0.06)	(0.06)	(0.06)	(0.12)	(0.13)	(0.13)
PAFS	-0.078***	-0.069***	-0.073***	-0.096***	-0.080**	-0.084***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Financial Risk	0.106**	-0.028	-0.038	0.402***	0.196**	0.174*
	(0.05)	(0.05)	(0.05)	(0.08)	(0.10)	(0.10)
North (dummy)		-0.116***	-0.387		-0.231***	0.080
		(0.04)	(0.35)		(0.08)	(0.65)
South (dummy)		0.133***	1.200***		0.075	1.662***
		(0.04)	(0.30)		(0.08)	(0.58)
AMUSA*North			0.073			-0.084
			(0.09)			(0.18)
AMUSA*South			-0.306***			-0.453***
			(0.08)			(0.16)
Constant	-3.394***	-3.546***	-3.914***	-3.646***	-3.539***	-4.288***
	(0.20)	(0.21)	(0.27)	(0.37)	(0.40)	(0.47)
N	2991	2991	2991	1639	1639	1639
adj R^2	0.008	0.018	0.027	0.038	0.046	0.049

Standard errors are in parentheses. All variables are in logs.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Robustness: UIR median and IQ range.

	(1)	(2)	(3)	(4)	(5)	(6)
	$MDN(c_f)$	$MDN(c_f)$	$MDN(c_f)$	$IQ(c_f)$	$IQ(c_f)$	$IQ(c_f)$
AMUSA (Φ_m)	-0.068**	-0.034	0.135*	-0.187***	-0.170**	0.090
	(0.03)	(0.04)	(0.07)	(0.07)	(0.07)	(0.12)
Rental price of land (ω_p)	0.016	0.137**	0.111*	-0.130	-0.125	-0.186
	(0.06)	(0.06)	(0.06)	(0.12)	(0.13)	(0.13)
North (dummy)		-0.069*	0.562*		-0.233***	0.186
		(0.04)	(0.34)		(0.07)	(0.62)
South (dummy)		0.104**	0.958***		-0.133*	1.704***
		(0.04)	(0.30)		(0.08)	(0.57)
AMUSA*North			-0.170*			-0.116
			(0.09)			(0.17)
AMUSA*South			-0.239***			-0.533***
			(0.08)			(0.16)
Constant	-2.856***	-3.222***	-3.794***	-2.460***	-2.377***	-3.197***
	(0.14)	(0.17)	(0.25)	(0.26)	(0.32)	(0.43)
N	3667	3667	3667	1960	1960	1960
adj R^2	0.001	0.006	0.007	0.007	0.012	0.017

Standard errors are in parentheses. All variables are in logs.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Robustness: regressions with AMUSA decomposed into accessibility and competition.

	(1) $AVG(c_f)$	(2) $AVG(c_f)$	(3) $SD(c_f)$	(4) $SD(c_f)$
Accessibility (Ψ_m)	-0.123***	-0.078**	-0.240***	-0.188***
	(0.03)	(0.04)	(0.07)	(0.07)
Total Sales Area (Θ_A)	0.082**	0.052	0.223***	0.199***
	(0.03)	(0.04)	(0.07)	(0.07)
Rental price of land (ω_p)	0.076	0.192***	-0.200*	-0.082
	(0.06)	(0.06)	(0.11)	(0.11)
North (dummy)		-0.095**		-0.261***
		(0.04)		(0.06)
South (dummy)		0.098**		0.009
		(0.04)		(0.07)
Constant	-2.627***	-3.040***	-2.357***	-2.737***
	(0.14)	(0.17)	(0.27)	(0.30)
N	3656	3656	1982	1982
adj R^2	0.008	0.014	0.014	0.025

Standard errors are in parentheses. All variables are in logs.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Robustness: excluding Rome and Milan (highest accessibility).

	(1) $AVG(c_f)$	(2) $AVG(c_f)$	(3) $AVG(c_f)$	(4) $SD(c_f)$	(5) $SD(c_f)$	(6) $SD(c_f)$
AMUSA (Φ_m)	-0.105***	-0.061*	0.064	-0.246***	-0.187***	-0.016
	(0.03)	(0.04)	(0.07)	(0.07)	(0.07)	(0.11)
Rental price of land (ω_p)	0.034	0.182***	0.153**	-0.207*	-0.084	-0.132
	(0.06)	(0.06)	(0.06)	(0.11)	(0.11)	(0.11)
North (dummy)		-0.122***	0.112		-0.249***	-0.216
		(0.04)	(0.33)		(0.06)	(0.56)
South (dummy)		0.100**	0.943***		0.012	1.433***
		(0.04)	(0.29)		(0.07)	(0.53)
AMUSA*North			-0.063			-0.011
			(0.09)			(0.15)
AMUSA*South			-0.241***			-0.415***
			(0.08)			(0.15)
Constant	-2.706***	-3.130***	-3.533***	-2.366***	-2.713***	-3.238***
	(0.14)	(0.16)	(0.24)	(0.26)	(0.30)	(0.37)
N	3654	3654	3654	1980	1980	1980
adj R^2	0.002	0.013	0.015	0.015	0.026	0.029

Standard errors are in parentheses. All variables are in logs.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Robustness: hypermarkets only.

	(1) $AVG(c_f)$	(2) $AVG(c_f)$	(3) $AVG(c_f)$	(4) $SD(c_f)$	(5) $SD(c_f)$	(6) $SD(c_f)$
AMUSA (Φ_m)	-0.105***	-0.061*	0.062	-0.245***	-0.186***	-0.014
	(0.03)	(0.04)	(0.07)	(0.07)	(0.07)	(0.11)
Rental price of land (ω_p)	0.034	0.181***	0.152**	-0.204*	-0.081	-0.130
	(0.06)	(0.06)	(0.06)	(0.11)	(0.11)	(0.11)
North (dummy)		-0.121***	0.106		-0.249***	-0.218
		(0.04)	(0.33)		(0.06)	(0.56)
South (dummy)		0.101***	0.938***		0.013	1.438***
		(0.04)	(0.29)		(0.07)	(0.53)
AMUSA*North			-0.062			-0.010
			(0.09)			(0.15)
AMUSA*South			-0.239***			-0.416***
			(0.08)			(0.15)
Constant	-2.705***	-3.129***	-3.527***	-2.375***	-2.722***	-3.246***
	(0.14)	(0.16)	(0.24)	(0.26)	(0.30)	(0.37)
N	3656	3656	3656	1982	1982	1982
adj R^2	0.002	0.013	0.014	0.015	0.026	0.029

Standard errors are in parentheses. All variables are in logs.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

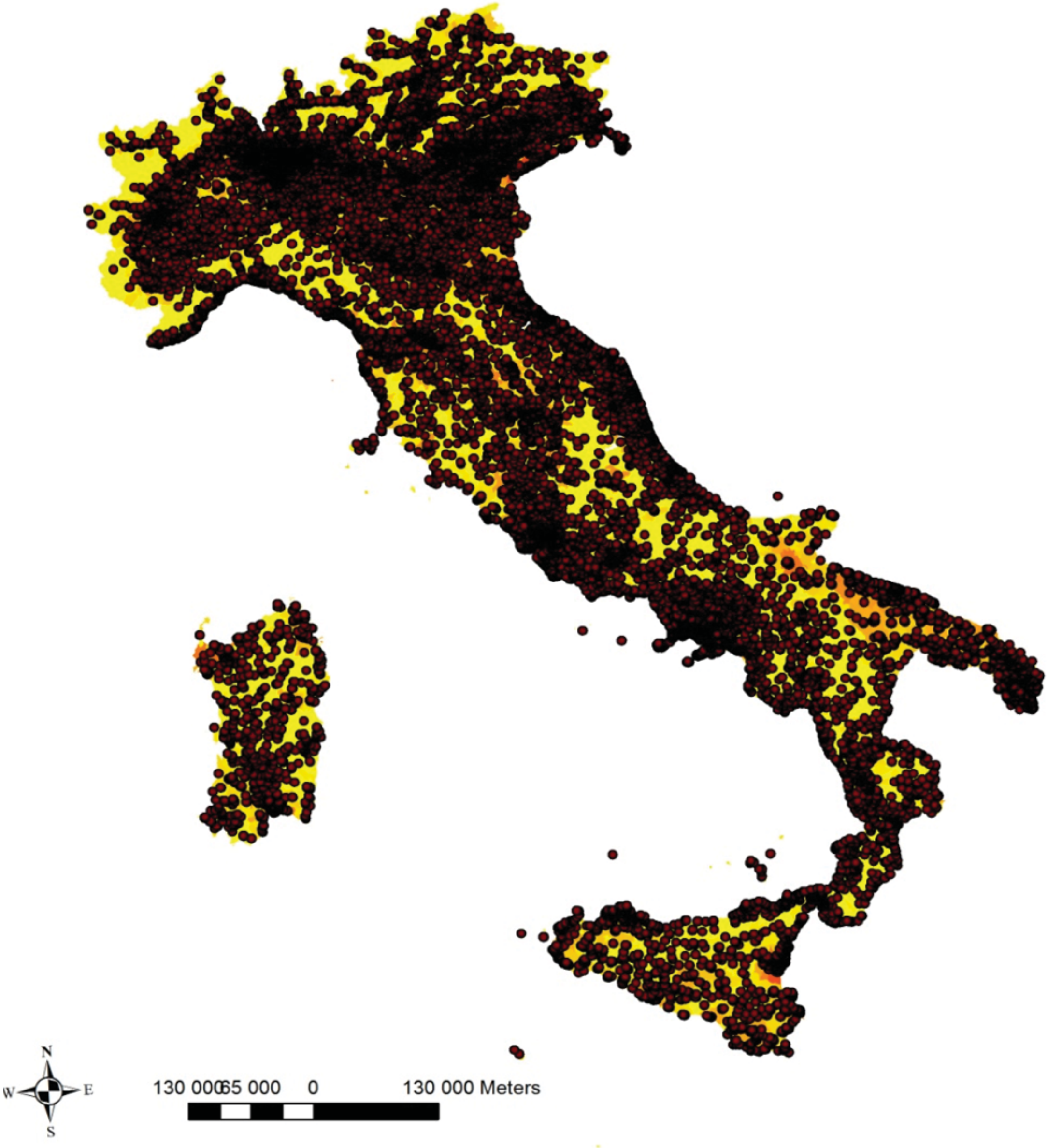

Figure 1: Geographic distribution of the stores in the Nielsen database.

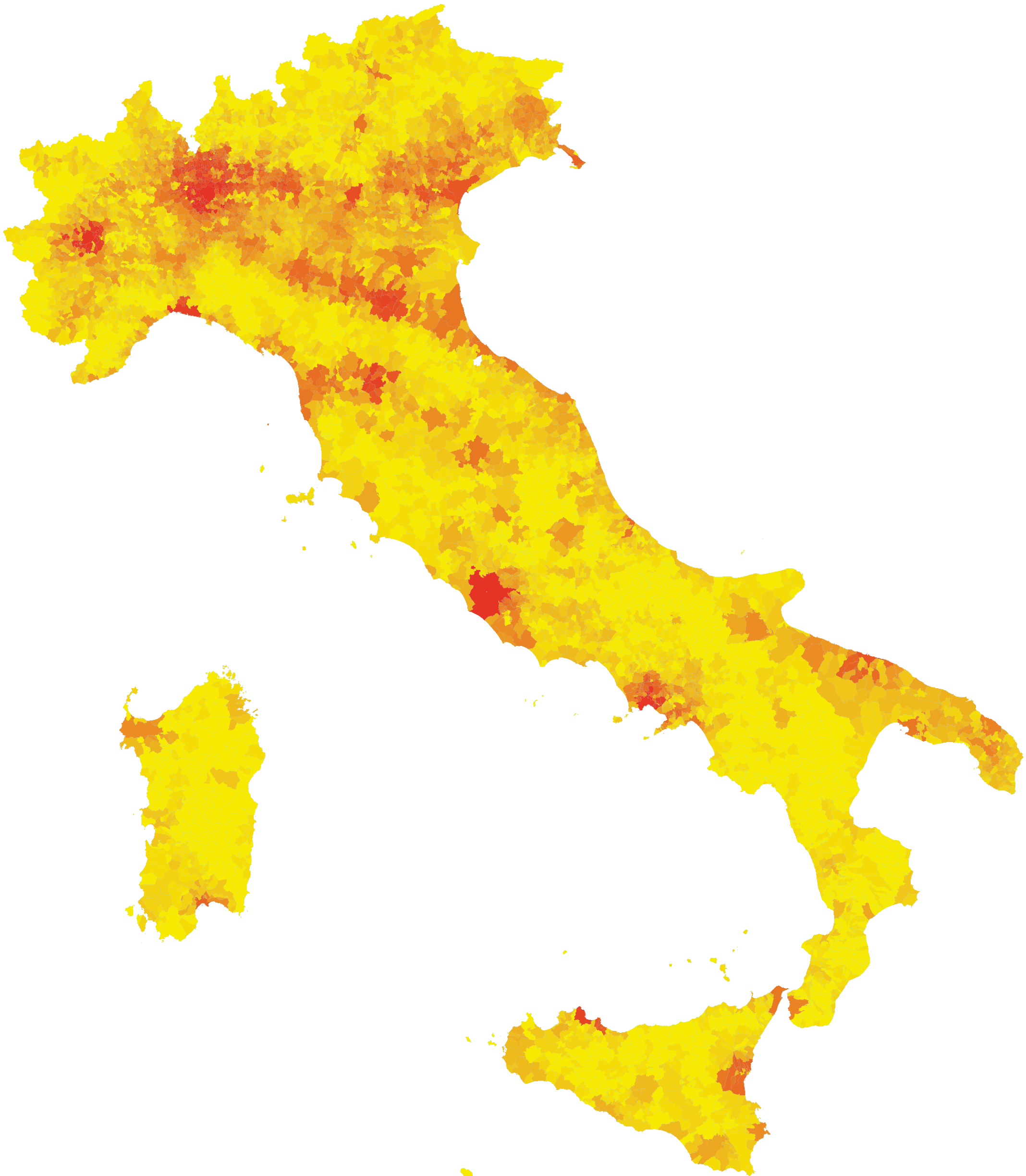

Figure 2: Potential Accessibility accessibility at the municipality level (Ψ_m).

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.it

17/10 *Sara Calligaris, Massimo Del Gatto, Fadi Hassan, Gianmarco I.P. Ottaviano, Fabiano Schivardi,* "The Productivity Puzzle and Misallocation: an Italian Perspective"
17/09 *Michele Battisti, Filippo Belloc. Massimo Del Gatto,* "Technology-specific Production Functions"
17/08 *Oliviero A. Carboni, Giuseppe Medda*, "Do Investment and Innovation Boost Export? An Analysis on European Firms"
17/07 *Edoardo Otranto, Massimo Mucciardi,* "Clustering Space-Time Series: A Flexible STAR Approach"
17/06 *Simone Franceschini, Gerardo Ettore Marletto,* "The dynamics of social capital during public participation: new knowledge from an on-going monitoring"
17/05 *Luca G. Deidda, Ettore Panetti,* "Banks' Liquidity Management and Systemic Risk"
17/04 *Luca Frigau, Tiziana Medda, Vittorio Pelligra,* "From the Field to the Lab An Experiment on the Representativeness of Standard Laboratory Subjects"
17/03 *William Addessi, Manuela Pulina,* "Sectoral Composition of Consumption Expenditure: A Regional Analysis"
17/02 *Claudio Detotto, Marta Meleddu, Marco Vannini,* "Cultural identity and willingness to protect and preserve art"
17/01 *Adriana Di Liberto, Laura Casula,* "Teacher assessments versus standardized tests: is acting "girly" an advantage?"
16/16 *Massimo Del Gatto, Carlo S. Mastinu,* "Sequentiality and Distance(s) in Cross-Border Mergers and Acquisitions: Evidence from Micro Data"
16/15 *Angelo Antoci, Simone Borghesi, Gianluca Iannucci*, "Green licenses and environmental corruption: a random matching model"
16/14 *Vittorio Pelligra, Tommaso Reggiani, Daniel John Zizzo,* "Responding to (un)reasonable requests"
16/13 *Pasqualina Arca Gianfranco Atzeni, Luca Deidda,* "Asset exemption in entrepreneurs' bankruptcy and the informative role of collateral"
16/12 *Miguel Casares, Luca Deidda, Jose E. Galdon-Sanchez*, "Loan production and monetary policy"
16/11 *Manuela Pulina, Valentina Santoni,* "An analysis on the Italian agricultural firms: effects of public subsidies"
16/10 *Tiziana Medda, Vittorio Pelligra, Tommaso Reggiani,* "Does Experience Affect Fairness, Reciprocity and Cooperation in Lab Experiments?"
16/09 *Bianca Biagi, Maria Giovanna Brandano, Manuela Pulina,* "The effect of tourism taxation: a synthetic control approach"
16/08 *Edoardo Otranto*, *Massimo Mucciardi* , "A Flexible Specification of Space–Time AutoRegressive Models"
16/07 *Leonardo Becchetti, Vittorio Pelligra, Francesco Salustri,* "Testing for Heterogeneity of Preferences in Randomized Experiments: A Satisfaction-Based Approach Applied to Multiplayer Prisoners' Dilemmas"
16/06 *Anna Bottasso, Maurizio Conti, Giovanni Sulis,* "Firm Dynamics and Employment Protection: Evidence from Sectoral Data"
16/05 *Oliviero A. Carboni, Giuseppe Medda,* "R&D, Export, and Investment Decision"

Finito di stampare nel mese di Dicembre 2017 Presso Centro Stampa
dell'Università degli Studi di Cagliari Via Università 40
09125 Cagliari

www.crenos.it

ISBN 978-88-9386-048-2

9 788893 860482 >