Spatially correlated preferences in international trade

Davide Arioldi[1]
June 2, 2021

ABSTRACT

Empirical evidences of extended gravity, spatial or sequential exporters and remote search of new trading partners have been theoretically justified by trade frictions and ad hoc dynamic models. We justify these empirical findings in a novel gravity model of trade, introducing spatially correlated consumer preferences in the Chaney (2008) model. Using the ratio of exports in a custom union, we are able to identify the spatial correlation parameter, through Monte Carlo Markov chain (HMC) method. Our results prove that consumers' preferences follow a spatially dependent structure, suggesting to take into account the spatial structure of demand in the gravity model of trade.

Keywords: spatially correlated preferences, trade network, extended gravity, spatial exporters.

JEL codes: C11, C21, F11, F14, F20, F40, F61

[1] Department of Economics, Università della Svizzera Italiana, Switzerland, Lugano (email: davide.arioldi@usi.ch).

1 Introduction

In recent years, the dynamic of firms' exports has received a great deal of attention. Das et al. (2007), modelling firms' exporting decision with sunk entry costs and plant level heterogeneity in export profits on a set of Colombian industries, find that entry costs are substantial and producers do not begin to export unless the present value of their expected future export profit stream is large. Furthermore, they state that history and expectation of producers are important determinants for the decision of being in a foreign market, even more than the value of the export profit that firms expect to earn in the current year[2]. Eaton et al. (2008), using transaction level data, observe that many Colombian firms enter foreign markets every year, selling small quantities to a single neighbor country, and almost half of them cease to export in the following year. The firms who survive expand their presence in the current destination and a sizeable fraction of them expands to other markets, depending on the initial foreign market. The empirical findings of Eaton et al. (2008), where many firms are jumping into and out of foreign markets, seem to be incompatible with large sunk costs, unless to suppose a two-tiered entry cost structure or serial correlated productivity of firm and product quality shocks. Moreover, Das et al. (2007) model does not explain the empirical sequential exporting findings described above. Nguyen (2012) and Albornoz et al. (2012) propose two new models to rationalize why firms wait to export and why many exporters fail. In the former model, demand is uncertain and imperfectly correlated among markets and firms choose to sequentially export in order to slowly learn about the possibility to succeed in new markets. The latter assumes that firms are uncertain about their export profitability but success factors are highly persistent over time and across destinations: therefore, entry in a foreign market allows firms to learn about their profit potentials in future and different markets. These new expectations are taken into account in firms' exporting decision and lead to a process of sequential exporting. Similar to the previous authors, Eaton et al. (2015) develop a search and learning model, where buyers reveal the appeal of the firms' product in a market, affecting the firms' propensity and cost to search for new clients. Chaney (2014), by modelling trade patterns as an international network, provides a further explanation for sequential entering. More specifically, firms export into markets where they have a contact, similar to social interactions (Jackson and Rogers, 2007). New contacts (trading partners) are searched both directly and indirectly; the formers using the existing network of contacts in the native market, the latter searching

[2] Firms prefer to continue to export in the foreign market even if their net profits are negative, because of the value they give to the possibility to export the next year without to pay entry costs.

remotely from the exporting markets. The predictions of Chaney's model are confirmed on a sample of French exporters, whose exports are geographically distributed conformed to the model.

More generally, standard gravity models do not completely capture the spatial correlation of trade, because predict trading patterns less spatially correlated than reality. Defever et al. (2015) provide robust causal evidence of extended gravity effects: using exports of a sample of Chinese firms after import liberalizations in US, EU and Canada, they prove that the probability of a firm to export in a country increases by about two percentage points for each additional prior export destination having a common border with the new country. Morales et al. (2017) quantify the impact of the extended gravity variables (common border, continent, language or similar income between new and previous foreign markets) on export entry costs, using a sample of Chilean firms. They find the sunk cost of entry in foreign markets is lower, from -19% to -38%, for markets having similarities with prior export destinations.

With this paper, we provide a framework to reconcile the extended gravity and sequential exporting findings with the traditional gravity model of trade. We extend the Chaney (2008) model of trade with heterogeneous firms by adding an unobserved country pairs and good specific preference parameter in the consumer's utility. As a result, we can shape trade flows as a function of the ratio between the consumers' preference parameter and fixed cost. Modelling the preference parameter or fixed cost as spatial dependent, the equation of trade internalizes the reinforced spatial pattern correlation. Sequential entering consequently emerges simply adding a time propagation effect to the fixed cost of entry or to the preference parameter, which are related to the geographical distribution of previous exports[3].

Differently from previous authors and supported by findings in empirical economics and marketing (Yang and Allenby, 2003; Rossi et al., 2005; Bradlow et al., 2005) we choose to spatially model the preference parameter. Spatial correlated preferences can explain many phenomena discovered by international trade scholars, such as the residual spatial correlation of traditional gravity model, the correlation over time and across destinations of export profitability and the "social network" effect discovered by Combes et al. (2005). Using a measure of social and business linkages inferred using migrations, Combes et al. (2005) find a positive impact of network linkages on inter-regional trade in France. Informational and social networks facilitate trade, overcoming informational barriers. Therefore, they posit that social interactions reduce fixed and variable costs to enter a foreign market. Garmendia et

[3] Sequential entering should emerges even only considering preferences correlated over time.

al. (2012), using Spanish data, confirm their results proving that the home bias disappears considering social and business networks. In our theoretical model with spatial correlated preferences, the network effect on firms' exports emerges formally, without assuming network lowers variable cost to export[4]. The preference parameter influences both the export value of each firm, the intensive margin, and the quantity of firms able to export, the extensive margin. According to our framework, network affects both the fixed cost to export, reducing informational barriers, and the preference parameter, boosting demand for imported and exported goods. Migrants, maintaining a network with their origin countries and with emigrants to other countries, can shape preferences, not only at idiosyncratic and bilateral country level but also on a world basis, promoting more homogenous preferences among all countries. In this paper, we do not assert that network does not affect exporting costs or that fixed costs to export are not spatially correlated[5], but we say that the geography of trade is widely affected by the spatial correlation of preferences. Including spatially correlated preferences in empirical studies and theoretical model, can advance the understanding of international phenomena and improve the evaluation of economic policies.

To the best of our knowledge, we are the first to provide general evidence of correlated preferences in the international trade structure and to propose a formal explanation of the extended gravity equation of trade. This paper is also related to the country of origin (COO) literature, started by Ditcher (1962) and widely explored by consumer's behaviours and marketing scholars in the last 40 years. Alternative recent research directions are proposed by Bertoletti et al. (2018), who develop a general equilibrium model of trade with non-homotetic indirectly additive preferences. Within their framework, both the extensive margin and intensive margin of trade depend positively on the per capita income of the destination country. Spatially enhanced effects can subsequently emerge because of spatially correlated income.

---

[4] In the classical model of trade with heterogeneous firm à la Melitz (2003), fixed cost of entry has no impact on the firm's export value but only on the extensive margin. It is therefore impossible to explain any increase of the intensive margin with a reduction of fixed costs (such as informational costs). Previous scholars have consequently assumed that networks of immigrants reduce variable exporting cost.

[5] The findings of Cavusgil and Zou (1994) and Artopoulus et al. (2011), documenting how product adaption and marketing strategy are key factors for firms export success, support this hypothesis. Product adaption, distribution chains, customers search and customer services can be model as fixed cost that increase with the distance to the target market. Moreover, knowledge of institutions and business practices are other fixed costs increasing with distance. Consequently, firms already exporting in foreign markets close to the target market could benefit from distribution chains, customer service supports and knowledge they have already developed in the previous markets. The fixed cost of entry in a new market is therefore increasing with the distance to the already reached markets.

The remainder of the paper is organized as follows: Chapter 2 presents reduced-form evidences that the spatial structure of export affects the probability to export to a new market (sequential entering) and the value exported. Chapter 3 provides a theoretical model with spatial correlated preferences able to explain the findings of Chapter 2. Chapter 4 proposes an identification strategy to estimate the idiosyncratic and correlation preference parameter on a subset of countries and goods, controlling for observable and unobserved fixed cost to export. Chapter 5 concludes.

2 Reduced form evidence

In this section, we provide reduced-form evidence that aggregate national export in a specific sector[6] and, more specifically, the spatial structure of the industry's trade network, affects both the probability to export to a market and the total traded value. The probability to sell goods to a foreign country is higher for sectors already exporting to markets close to the foreign destination and it increases with the value exported to those markets; moreover, this probability is higher if the foreign destination and the prior markets have a trading relationship in the same sector. Similar results emerge for the national value of exports: the more a sector exports to countries close to the target market, the higher is the value exported to the target market. These effects are robust, even after controlling for the extended gravity variables, which should be a proxy for correlated fixed costs to export, according to Morales et al. (2017).

Data source – We use product level data aggregate at the 2 digit-level of the Standard International Trade Classification, Revision 2, over the period 1980-2000. The data comes from the same source of Feenstra et al. (2005) and includes trade from 155 exporters to 154 importers, accounting for about 98% of the world trade. We add zero trade flows to this data for every combination of exporter, importer and sector that is not reported. Our final dataset includes about 33 million observations. Every product is exported on average to 12 different countries, with a minimum average of 0.4 exporting markets for the SITC 35 class (fish, dried, salted or in brine; smoked fish) and a maximum average of 24 exporting markets for the SITC 65 Class (Textile yarn, fabrics, made-up articles and related products). About 30% of the national industries[7] in our sample have never exported to any country and only 3 classes of products (specialized industrial machinery, road vehicles, medicinal and pharmaceutical products) have been exported to at least 150 countries. In addition to the data on trade flows, we add

[6] We assume that each product *h* is produced by sector *h*.
[7] We assume that for each country-SITC product class corresponds a set of firms belonging to the same nation and producing the specific good.

geographical variables (such as the population weighted distance, contiguity and binary variables for regional trade agreement, common or former colony, same language), economic variables (such as gross domestic product) and extended gravity variables[8]. This final dataset includes about 12 million observations, with 2,147,868 positive flows.

Regression specifications – We estimate probit and linear regressions (OLS) with different specifications of the remote distance variable. Our dependent variables are the exporting status in the target market j for good h produced in country i in year t + 1 and the trade value from country i to country j for product h in year t. Using two set of countries K (K_0 and K_1) we compute two variables for the distance between the other exporting destinations K and the target market j, modelling them as the log of averaged distances (population weighted) from j to K, and two variables for the total exports of product h from country i to countries K_0 and K_1. Countries belonging to K_0 are all the countries, different from j, where country i sell the good h while countries included in K_1 are a subset of K_0, whenever the countries are already exporting the good h to the consumers located in the country j (that are countries jointly belonging to the set of importers of j and to the set of exporters of i, for sector h).

With the probit specification, we also test if belonging to different trade networks (as above, at the product level h) affects the probability to directly export to the target market using the minimum distance path between sector h of country i and the target market j. The minimum distance path is defined as the minimum number of exporting markets the producing sector h has to pass through to reach the target market j[9]. We include in our model a set of two binary variable for the path distance: the first ($1[\min(pathdist^h_{i,j,t}) = \infty]$) is equal to one if countries i and j do not belong to the same trade network (the path distance is equal to infinite) while the second ($1[\min(pathdist^h_{i,j,t}) = 1]$) is set to 1 if country i and j belong to the same trade network and country i exports to at least one country k belonging to the set of importer of j. The base category for this variable is therefore the path distance different to infinite and one.

---

[8] We compute extended gravity variables (extended contiguity, common language, common colony, common currency, common religion and common legal system) as in Morales et al. (2017), Albornoz et al. (2012) or Defever et al. (2015) to control for potential different fixed costs to export. Geographical variables are from CEPII while economic variables from Penn World table.

[9] If sector h in country i exports directly to country j, the minimum distance path between i and j for sector h is zero; if sector h does not export directly to j, but it exports in k and industry h in k exports directly to j, the minimum distance path is equal to 1.

As control variables, we add the difference between the per capita GDP of the two countries, to control for cost that firms face to adapt the production chain to the quality requested in the new market[10], and the number of exporting markets of sector *h*, in order to control for the experience and propensity to export to foreign markets. We even control for all the set of traditional gravity variables, as the log of the population weighted distance between the two countries, the log of the GDP of countries *i* and *j* and a set of binary variables controlling for Regional Trade Agreement, contiguity, common language, common colonizer, common currency, GATT/WTO membership for exporter and importer and the share of population with common religion. For the probit model we add the exporting status of country *i* in market *j* for the product *h* at time *t*, that controls for the resilience of the exporting status, and the import growth of country *j* for product *h*, as in Chaney (2014). As a robustness test, we include a set of extended gravity variables (as in Albornoz et al., 2012, or Defever et al., 2015) for contiguity, common language, common colony/colonizer, common currency, common religion and common legal system. These variables are equal to 1 if product *h* produced by country *i* is exported to at least one country *k* sharing some characteristics with country *j*[11]. These variables can control for fixed (even sunk) costs to export which are correlated among destinations, as in Morales at al. (2017).

We estimate different specifications of the probit model in Equation 1 and OLS in Equation 2. In Equation 1, we chose to use, as dependent variable, the lead exporting status (at *t+1*) instead of the exporting status at time *t*, in order to give results comparable to the previous literature[12].

$$\Pr\left(export_{i,j,t+1}^{h} > 0 \middle| observables\right) = \Phi\left(\delta 1\left[export_{i,j,t}^{h} > 0\right] + \gamma_1 \ln(avg\_dist_{K,j,t}^{h}) + \gamma_2 ln\left(\sum_k export_{i,k,t}^{h}\right) + \beta Controls_{i,j,t} + Exporter_i + Importer_j + Year_t + Extended\ Gravity\ Variables_{i,j,t}^{h}\right) \quad (1)$$

$$\ln(export_{i,j,t}^{h}) = \gamma_3 \ln\left(avg\_dist_{K,j,t}^{h}\right) + \gamma_4 ln\left(\sum_k export_{i,k,t}^{h}\right) + \beta Controls_{i,j,t} + Exporter_i + Importer_j + Year_t + Extended\ Gravity\ Variables_{i,j,t}^{h} + \varepsilon_{i,j,t}^{h} \quad (2)$$

[10] Murphy and Shleifer (1997) have shown that countries tend to trade with partners with similar level of development, producing similar quality products.

[11] For the sake of clarification, extended contiguity is equal to one if country *i* exports good *h* to at least one country *k* with a common border with *j*; equally, extended common currency is equal to 1 if country *i* exports good *h* to at least one country *k* having the same currency of country *j* and so on.

[12] Performing our analysis with export status at time *t* does not significantly change the results and all the conclusions remain meaningful. The same occurs when estimating Equation 2 with the lag of the averaged remote distances and lagged total exports.

Given that we are assuming spatially correlated preferences, we expect to find negative values for the coefficients of remote distance (parameters γ_2 and γ_4), because industries exporting product *h* to countries close to *j* are expected to be more likely to export and to sell more goods to market *j*. γ_1 and γ_3 are instead supposed to be positive, given that the more the product *h* is exported to the other markets *K*, the more likely it will be exported to the market *j*, with larger quantities (remote quantity).

We even propose a different specification of remote distance (γ_2^* and γ_4^*) and remote quantity (γ_1^* and γ_3^*), where the values are computed using the subset of countries that import from the producing country and export to the target market.

Results – Estimates of Equations 1 and 2 are reported in Table 1 and 2, respectively.

Table 1 – Export Network Effect on the probability to export to market j

Dependent variable $export^h_{i,j,t+1} > 0$	Par	(1)	(2)	(3)	(4)	(5)	(6)
$1[export^h_{i,j,t} > 0]$	$\delta 1$	2.102*** (0.0453)	1.995*** (0.0435)	1.842*** (0.0490)	1.816*** (0.0471)	1.899*** (0.0430)	1.841*** (0.0471)
$\sum_k 1[export^h_{i,k,t} > 0]$	β_1	0.0171*** (0.00046)	0.0151*** (0.00044)	0.01*** (0.00037)	0.0096*** (0.00038)	0.0099*** (0.00037)	0.0085*** (0.00036)
$ln(GDPpc_{i,t} - GDPpc_{j,t})^2$	β_1	-0.00505* (0.00210)	-0.006** (0.00199)	-0.0055** (0.00209)	-0.0062** (0.00202)	-0.0062** (0.00208)	-0.0059** (0.00212)
$lnDist_{i,j}$	β_1	-0.394*** (0.0116)	-0.376*** (0.0113)	-0.356*** (0.0111)	-0.348*** (0.0109)	-0.387*** (0.0113)	-0.358*** (0.0111)
$1[\min(pathdist^h_{i,j,t}) = 1]$	δ_1	0.258*** (0.0234)	0.188*** (0.0240)	0.0358 (0.0301)	0.0260 (0.0296)	0.109*** (0.0243)	0.0423 (0.0278)
$1[\min(pathdist^h_{i,j,t}) = \infty]$	δ_2	-0.105+ (0.0541)	-0.140* (0.0573)	-0.219** (0.0669)	-0.222*** (0.0666)	-0.176** (0.0650)	-0.214** (0.0681)
$\ln \sum_k 1[\mathrm{export}^h_{k,j,t} > 0]\mathrm{export}^h_{i,k,t}$	γ_1^*			0.312*** (0.00967)	0.269*** (0.00947)		0.179*** (0.00912)
$\ln \frac{\sum_k 1[\mathrm{export}^h_{i,k,t} > 0 \,\&\, \mathrm{export}^h_{k,j,t} > 0](\mathrm{Dist}^h_{k,j,t})}{\sum_k 1[\mathrm{export}^h_{i,k,t} > 0 \,\&\, \mathrm{export}^h_{k,j,t} > 0]}$	γ_2^*			-0.568*** (0.0150)	-0.482*** (0.0138)		-0.440*** (0.0140)
$\ln \sum_k \mathrm{export}^h_{i,k,t}$	γ_1					0.111*** (0.00423)	0.0873*** (0.00395)
$\ln \frac{\sum_k 1[\mathrm{export}^h_{i,k,t} > 0](\mathrm{Dist}^h_{k,j,t})}{\sum_k 1[\mathrm{export}^h_{i,k,t} > 0]}$	γ_2					-0.0159** (0.00536)	
Extended Contiguity			0.231*** (0.00556)		0.193*** (0.00484)	0.223*** (0.00563)	
Extended Common Language			0.105*** (0.00457)		0.0525*** (0.00404)	0.0758*** (0.00407)	
Extended Common Colony/Colonizer			0.130*** (0.00590)		0.0895*** (0.00534)	0.120*** (0.00576)	
Extended Common Currency			-0.00187 (0.00440)		-0.00439 (0.00435)	-0.000271 (0.00437)	
Extended Common Religion			-0.0657 (0.0556)		-0.102+ (0.0532)	-0.0993* (0.0481)	
Extended Common Legal System			0.165*** (0.00702)		0.0644*** (0.00546)	0.0959*** (0.00643)	
Control variables		Y	Y	Y	Y	Y	Y
Exporter Fixed Effect		Y	Y	Y	Y	Y	Y
Importer Fixed Effect		Y	Y	Y	Y	Y	Y
Year Fixed Effect		Y	Y	Y	Y	Y	Y
N		11564608	11564608	11564608	11564608	11564608	11564608

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Notes: This table shows the coefficients of the Probit estimation of Equation 1 for 62 products (SITC rev. 2 at 2 digit) traded between 155 countries from 1980 to 2000. The dependent variable is equal to 1 if the product h produced by industry h in country i is exported to country j at time t + 1. Control variables include import growth of country j, the log of GDP of the two countries and a set of binary variables for regional trade agreement (as reported by WTO), contiguity, common language, common colonizer, common currency, importer GATT membership, exporter GATT membership and the share of population with same religion. Standard errors, clustered at the product level, are in parenthesis.

Table 1 shows the results of the Probit estimation for different specifications of Equation 1. The coefficients γ_1 (the level of export to other countries) and γ_2 (the remote distance) have the expected signs and are significative at the 0.1 percent level. These results confirm that we are more likely to observe exports of product *h* from country *i* to country *j* if country *i* already exports the same good *h* to countries close to j and that the probability to export to market j is positively correlated with the total value of good h exported to the other countries close to j. Furthermore, as shown by the coefficient of the path distance variable δ_2, the probability to export to j is higher if the product h is exported to at least one country k belonging to the trade network of j. This effect is robust even considering the full set of extended gravity variables that should control for differences in fixed costs to export and potentially capturing, at least partially, preferences' similarities. Having exported in the previous year to countries sharing a border or having a common language, legal system or the same colonizer of the target market increases the probability to export to the target market. Extended common currency and common religion seem however having a null or a negative impact[13]. With reference to the other control variables, all the signs and significance levels are as expected. The probability to export to the target market is positively correlated with the number of previous foreign destinations while decreases with geographic and GDP distances.

We posit that even the value of trade is influenced by the structure of the trade network, specifically by the total exporting value to the other countries K and by their distances to market j. We assume a complex ("*satellite*") gravitational effect, where distance exhibits both direct (from the local market) and indirect (from the remote markets) effect. We estimate Equation 2 as a standard empirical gravity equation, with exporter, importer, year and sector fixed effects, using the same group of control variables from Equation 1, except for the import growth variable, the path distance variables and the exporting status variable. As standard in all gravity equations, our dependent variable is the log of the aggregated trading value of product h from country i to country j in year t. The impact of the export structure on the value of trade is reported on Table 2, parameters γ_1 and γ_2.

[13] The negative impact of the extended common legal system variable in Models 6 seems to be related to the correlation with the value of export to countries K. Extended results are available upon request.

Table 2 – Export Network Effect on the value of trade (intensive margin)

Dependent variable $\ln export^{h}_{i,j,t}$		(1)	(2)	(3)	(4)	(5)	(6)
$\sum_{k} 1[export^{h}_{i,k,t} > 0]$	β_1	0.0408*** (0.00106)	0.0397*** (0.00104)	0.0311*** (0.00140)	0.0312*** (0.00141)	0.015*** (0.00093)	0.015*** (0.00095)
$ln(GDPpc_{i,t} - GDPpc_{j,t})^2$	β_2	-0.04*** (0.00540)	-0.041*** (0.00534)	-0.038*** (0.00530)	-0.038*** (0.00528)	-0.042*** (0.00521)	-0.042*** (0.00517)
$lnDist_{i,j}$	β_3	-0.806*** (0.0334)	-0.791*** (0.0333)	-0.787*** (0.0341)	-0.781*** (0.0342)	-0.822*** (0.0313)	-0.814*** (0.0314)
$\ln \sum_{k} 1[\mathrm{export}^{h}_{k,j,t} > 0]\mathrm{export}^{h}_{i,k,t}$	γ_3^*			0.557*** (0.0659)	0.513*** (0.0640)		
$\ln \frac{\sum_k 1[\mathrm{export}^{h}_{i,k,t} > 0 \,\&\, \mathrm{export}^{h}_{k,j,t} > 0](\mathrm{Dist}^{h}_{k,j,t})}{\sum_k 1[\mathrm{export}^{h}_{i,k,t} > 0 \,\&\, \mathrm{export}^{h}_{k,j,t} > 0]}$	γ_4^*			-0.864*** (0.103)	-0.779*** (0.0996)		
$\ln \sum_{k} \mathrm{export}^{h}_{i,k,t}$	γ_3					0.632*** (0.0217)	0.628*** (0.0219)
$\ln \frac{\sum_k 1[\mathrm{export}^{h}_{i,k,t} > 0](\mathrm{Dist}^{h}_{k,j,t})}{\sum_k 1[\mathrm{export}^{h}_{i,k,t} > 0]}$	γ_4					-0.199*** (0.0126)	-0.202*** (0.0125)
Extended Contiguity			0.200*** (0.0156)		0.141*** (0.0115)		0.124*** (0.0128)
Extended Common Language			0.216*** (0.0232)		0.134*** (0.0163)		0.0879*** (0.0181)
Extended Common Colony/Colonizer			0.0209 (0.0171)		-0.0248 (0.0162)		-0.0181 (0.0161)
Extended Common Currency			0.0176* (0.00842)		0.0123 (0.00848)		0.00815 (0.00859)
Extended Common Religion			-0.657** (0.207)		-0.676** (0.202)		-0.707*** (0.154)
Extended Common Legal System			0.0623** (0.0182)		0.000091 (0.0180)		-0.066*** (0.0172)
Control variables		Y	Y	Y	Y	Y	Y
Exporter Fixed Effect		Y	Y	Y	Y	Y	Y
Importer Fixed Effect		Y	Y	Y	Y	Y	Y
Year Fixed Effect		Y	Y	Y	Y	Y	Y
N		2147868	2147868	2147868	2147868	2147868	2147868

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Notes: This table shows the coefficients for the OLS estimation of a gravity equation of trade, for 62 products (SITC rev. 2- 2 digit, level) between 155 countries from 1980 to 2000. The dependent variable is the log of trade value for the product h produced in country i and exported to j at time t. Control variables include the log of GDP of the two countries and a set of binary variables for regional trade agreement (as reported by WTO), contiguity, common language, common colonizer, common currency, importer GATT membership, exporter GATT membership and the share of population with same religion. Standard errors clustered at the product level are in parenthesis.

The closer is market *j* to the other exporting destinations *K* and the more country *i* sells good *h* to these countries *K*, the higher market *j* imports the good *h* produced by country *i*. As shown in Table 1 and 2, industries not only are more likely to export in countries close to their other destinations, but also they export more.

3 Theoretical model

Our previous empirical findings support the idea of enhanced spatial effects and path dependence structure of trade. As shown by other scholars, these spatial effects are not completely caught by traditional gravity models. With our reduced form function, we are not able to identify if these evidences originate from spatially (or network) correlated preferences of consumers (demand side) or from spatially correlated fixed costs to access foreign markets (supply side). To define formally our framework, we extend the Chaney (2008) model of trade with heterogeneous firms by adding an exogenous preference parameter in the consumer's utility. Supported by empirical findings in the Country of Origin (COO) and marketing literature, we assume consumers' preferences are country pairs, goods specific and correlated among consumers[14]. As in Chaney (2008), we consider *N* countries with population L_n. Each firm produces one differentiated good using only labor, with a given productivity ϕ. Consumers in each country *n* consume $q_h(\omega)$ units of each variety ω of good *h* and q_0 units of good 0, the homogeneous good. Sectors *H* produce a continuum of differentiated goods. The utility of consumer is:

$$U = q_0^{\mu_0} \left[\prod_{h=1}^{H} \int_{\Omega_h} (\alpha_{ij}^h)^{\frac{1}{\sigma_h}} q_h(\omega)^{\frac{\sigma_h - 1}{\sigma_h}} d\omega \right]^{\frac{\sigma_h}{\sigma_h - 1}\mu_h} \tag{3}$$

where $\mu_0 + \sum_1^H \mu_h = 1$. α_{ij}^h is the exogenous preference shifter of consumers in country *j* for the goods of sector *h* produced in country *i*, as in Feenstra et al. (2018)[15], while $\sigma_h > 1$ is the elasticity of substitution between two varieties of good *h*. The homogeneous good 0 is used as numeraire and is produced with unitary constant returns to scale[16]. It is freely traded with price equal to 1, so that if country *n* produces the good, the wage in the country is w_n. As standard in literature, we assume that each country produces the numeraire, to simplify the analysis[17]. For a firm in country *i*, the cost of producing and selling *q* unit of good *h* to country *j* is:

$$c_{ij}^h(q) = \frac{w_i \tau_{ij}^h}{\varphi} q + f_{ij}^h \tag{4}$$

[14] The willingness to pay for Italian or Spanish hams is higher in Europe but not in middle-east Muslim countries. German or Japan cars are worldwide perceived as high quality products, or European consumers prefer Swiss watches to Japan watches.
[15] $\sum_{i \neq j} \alpha_{ij}^h = 1$
[16] One unit of labor in country n produces w_n units of good 0.
[17] As Chaney (2008) specifies, the assumption hold as long as μ_0 or trade barriers are large enough.

where $\tau_{ij}^{h} > 1$ is the variable trade cost in the form of an "iceberg" transportation cost and f_{ij}^{h} is the fixed cost to export from country *i* to country *j*. All countries have access to the same technology and, given the presence of fixed cost, firms produce under increasing returns to scale. The unit labor productivity φ is drawn by each firm from a Pareto distribution function, as in Helpman, Melitz and Yeaple (2004) with shape parameter $\gamma_h > \sigma_h - 1$. The distribution of productivity is

$$P(\tilde{\varphi}_h < \varphi) = G_h(\varphi) = 1 - \varphi^{-\gamma_h} \qquad (5)$$

distributed over $[1, \infty)$. Higher value of γ_h implies that firms' productivities are more homogeneous and concentrated among the lower part of the distribution. The condition $\gamma_h > \sigma_h - 1$ ensures that the size distribution of firms has a finite mean in equilibrium. Similarly to Chaney (2008), we assume the total mass of potential entrants in country *n* proportional to $w_n L_n$ in order to simplify the analysis. Hence, bigger and wealthier countries have more potential entrants. Furthermore, each worker owns own shares of a global fund collecting profits from all the firms and redistributing them in units of the numeraire good to the shareholders. Total expenditure Y_j of workers in country *j* is therefore the total income, $(1+\pi)w_j L_j$, where π is the dividend per share of the global mutual fund.

Firms are price setters and given that the demand function is isoelastic, the optimal price is a constant mark-up over the unit cost, including the iceberg transportation cost. The demand for exports from country *i* to country *j* in sector *h*, faced by a firm with productivity φ, is

$$x_{ij}^{h}(\varphi) = p_{ij}^{h}(\varphi) q_{ij}^{h}(\varphi) = \mu_h Y_j \, \alpha_{ij}^{h} \left(\frac{p_{ij}^{h}(\varphi)}{P_j^{h}} \right)^{1-\sigma_h} \qquad (6)$$

where $P_j^h = \left[\sum_k^N \alpha_{kj}^h p_{kj}^{h\;1-\sigma_h} \right]^{\frac{1}{1-\sigma_h}}$ is the price index for good *h* in country *j*. Because firms face a fixed cost to export, f_{kj}^h, only the firms with productivity φ above the threshold $\tilde{\varphi}_{kj}^h$ can export to country *j*. Given the assumption of exogenous entry (proportional to $w_k L_k$), the price index P_j^h is defined as

$$P_j^h = \left[\sum_{k=1}^{N} w_k L_k \int_{\tilde{\varphi}_{kj}^h}^{\infty} \alpha_{kj}^h \left(\frac{\sigma_h}{\sigma_h - 1} \frac{w_k \tau_{kj}^h}{\varphi} \right)^{1-\sigma_h} dG_h(\varphi) \right]^{\frac{1}{1-\sigma_h}} \qquad (7)$$

and the net profit that a firm with productivity φ, producing good *h* in country *k*, earns exporting to country *l* is

$$\pi_{kl}^{h}\,(\varphi) = [p_{kl}^{h}\,(\varphi) - c_{kl}^{h}\,(\varphi)]q_{kl}^{h}\,(\varphi) - f_{kl}^{h} \tag{8}$$

Aggregate profit are therefore as in Chaney (2008)

$$\pi = \frac{\sum_{h=1}^{H}\sum_{k,l=1}^{N} w_k L_k \left(\int_{\tilde{\varphi}_{kl}^{h}}^{\infty} \pi_{kl}^{h}\,(\varphi) dG_h(\varphi)\right)}{\sum_{n=1}^{N} w_n L_n} \tag{9}$$

To compute the general equilibrium solution of this system we have to specify the cut-off productivity level $\tilde{\varphi}_{kj}^{h}$ above which firms export to country *j*. Plugging the demand of consumers and the price settled by firms in the profit equation we have the following equation for net profit $\pi_{ij}(\varphi)$

$$\pi_{ij}(\varphi) = \frac{\mu_h}{\sigma_h} Y_j\, \alpha_{ij}^{h} \left(\frac{\sigma_h}{\sigma_h - 1}\, \frac{w_i \tau_{ij}^{h}}{\varphi P_j^{h}}\right)^{1-\sigma_h} - f_{ij}^{h} \tag{10}$$

Defining the threshold $\tilde{\varphi}_{ij}^{h}$ as the level of productivity where the profit of the firm with productivity φ producing good *h* in country *i* and exporting in country *j* is null ($\pi_{ij}(\varphi) = 0$), we can rearrange the previous equation as

$$\tilde{\varphi}_{ij}^{h} = \left(\frac{\sigma_h}{\mu_h}\right)^{1/(\sigma_h - 1)} \frac{\sigma_h}{\sigma_h - 1} \left(\frac{f_{ij}^{h}}{Y_j \alpha_{ij}^{h}}\right)^{1/(\sigma_h - 1)} \frac{w_i \tau_{ij}^{h}}{P_j^{h}} \tag{11}$$

$\tilde{\varphi}_{ij}^{h}$ is therefore the productivity level below which any firms does not export to country *j*.

We observe that the preference parameter, α_{ij}^{h}, has a balancing effect with respect to the fixed cost, f_{ij}^{h}. A proportional increase of preferences for good *h* produced in country *i* is equivalent to a proportional reduction in the fixed cost of export f_{ij}^{h}. Formally, $\frac{\partial\, ln\, \tilde{\varphi}_{ij}^{h}}{\partial\, ln\, \alpha_{ij}^{h}} = -\frac{\partial\, ln\, \tilde{\varphi}_{ij}^{h}}{\partial\, ln\, f_{ij}^{h}} = 1/(\sigma_h - 1)$.

Thanks to the assumptions that wages are exogenously pinned down in the homogeneous sector 0 and entrants are exogenously determined, the equilibrium price index is given by the solution of the following system of equations:

$$\begin{cases} P_j^h = \left[\sum_{k=1}^{N} w_k L_k \int_{\tilde{\varphi}_{kj}^h}^{\infty} \alpha_{kj}^h \left(\frac{\sigma_h}{\sigma_h - 1} \frac{w_k \tau_{kj}^h}{\varphi}\right)^{1-\sigma_h} dG_h(\varphi)\right]^{\frac{1}{1-\sigma_h}} \\ \tilde{\varphi}_{kj}^h = \left(\frac{\sigma_h}{\mu_h}\right)^{1/(\sigma_h-1)} \frac{\sigma_h}{\sigma_h - 1}\left(\frac{f_{kj}^h}{Y_j \alpha_{kj}^h}\right)^{1/(\sigma_h-1)} \frac{w_k \tau_{kj}^h}{P_j^h} \end{cases} \quad (12)$$

Considering that the distribution of $G_h(\varphi)$ is a Pareto with shape parameter γ_h, we can solve for the integral and rearranging the equation expressing P_j^h as

$$P_j^h = \Phi_1 \, Y_j^{\frac{1}{\gamma_h} - \frac{1}{\sigma_h - 1}} \theta_j^h \quad (13)$$

where $\theta_j^h = \left(\frac{Y^*}{Y_k}\right)^{\frac{1}{\gamma_h}} \sum_{k=1}^{N} (w_k \tau_{kj}^h) f_{kj}^{h\,-\frac{1}{\gamma_h}} \left(\frac{f_{kj}^h}{\alpha_{kj}^h}\right)^{\frac{1}{\sigma_h - 1}}$ and $\Phi_1 = \frac{\sigma_h}{\sigma_h - 1}\left(\frac{\gamma_h - (\sigma_h - 1)}{\gamma_h}\right)^{\frac{1}{\gamma_h}} \left(\frac{\sigma_h}{\mu_h}\right)^{\left(\frac{1}{\sigma_h - 1} - \frac{1}{\gamma_h}\right)} \left(\frac{1+\pi^*}{Y^*}\right)^{\frac{1}{\gamma_h}}$ with Y* and π^* being, respectively, world output and profit computed in equilibrium.

By plugging the price index P_j^h into the demand function and the productivity threshold we can now compute the value of exports of an individual firm with labour productivity φ. As in Chaney (2008), we can simultaneously solve the system for firm level export, productivity threshold and total world profit.

The solution is given by plugging P_j^h into the following system of equations:

$$
\begin{cases}
x_{ij}(\varphi) = \begin{cases} \mu_h \left(\dfrac{\sigma_h}{\sigma_h - 1}\right)^{1-\sigma_h} Y_j\, \alpha_{ij}^h \left(\dfrac{w_i \tau_{ij}^h}{P_j^h}\right)^{1-\sigma_h} \varphi^{\sigma_h - 1} & if\ \varphi \geq \tilde{\varphi}_{ij}^h \\ 0 & if\ \varphi < \tilde{\varphi}_{ij}^h \end{cases} \\
\tilde{\varphi}_{ij}^h = \left(\dfrac{\sigma_h}{\mu_h}\right)^{1/(\sigma_h - 1)} \dfrac{\sigma_h}{\sigma_h - 1} \left(\dfrac{f_{ij}^h}{Y_j \alpha_{ij}^h}\right)^{1/(\sigma_h - 1)} \dfrac{w_i \tau_{ij}^h}{P_j^h} \\
Y_i = (1+\pi) w_i L_i \\
\pi = \sum_{h=1}^{H} \pi_h
\end{cases}
\tag{14}
$$

As a result, we can define the export of product *h* to country *j* for a firm producing in country *i* as:

$$
x_{ij}(\varphi) = \begin{cases} \Phi_2 \left(\dfrac{Y_j}{Y}\right)^{\frac{\sigma_h - 1}{\gamma_h}} \alpha_{ij}^h \left(\dfrac{w_i \tau_{ij}^h}{\theta_j^h}\right)^{1-\sigma_h} \varphi^{\sigma_h - 1} & if\ \varphi \geq \tilde{\varphi}_{ij}^h \\ 0 & if\ \varphi < \tilde{\varphi}_{ij}^h \end{cases} \tag{15}
$$

With threshold $\tilde{\varphi}_{ij}^h = \Phi_3 \left(\frac{Y}{Y_j}\right)^{\frac{1}{\gamma_h}} \frac{w_i \tau_{ij}^h}{\theta_j^h} \left(\frac{f_{ij}^h}{\alpha_{ij}^h}\right)^{1/(\sigma_h - 1)}$, $\Phi_2 = \sigma_h \Phi_3^{1-\sigma_h}$, $\Phi_3 = \left[\frac{\sigma_h}{\mu_h} \frac{\gamma_h}{\gamma_h - (\sigma_h - 1)} \frac{1}{1+\pi}\right]^{\frac{1}{\gamma_h}}$ and π computed as in Chaney (2008).

The preference parameter of the country of origin has therefore an effect on the intensive and extensive margins of trade, lowering the productivity cutoff in the latter case[18].

In order to compute the value of aggregate trade from country *i* to country *j* for good *h* (X_{ij}^h) ,we have to sum the exporting value to country *j* of all firms in country *i* producing the good *h*, having a productivity at least equal to the productivity threshold $\tilde{\varphi}_{ij}^h$. With the assumption of exogenous potential entry, we define total exports as $X_{ij}^h = w_i L_i \int_{\tilde{\varphi}_{ij}^h}^{\infty} x_{ij}^h(\varphi) dG_h(\varphi)$

[18] A lower cutoff implies that a higher number of firms can access the foreign market *j*.

Plugging Equation 15 and the corresponding threshold $\tilde{\varphi}_{ij}^{h}$ into X_{ij}^{h} and considering $Y_i = (1 + \pi)w_i L_i$ we can derive total exports[19] as:

$$X_{ij}^{h} = \mu_h \frac{Y_i\, Y_j}{Y} \left(\frac{w_i \tau_{ij}^{h}}{\theta_j^{h}}\right)^{-\gamma_h} \alpha_{ij}^{h} \left(\frac{\alpha_{ij}^{h}}{f_{ij}^{h}}\right)^{\frac{\gamma_h}{(\sigma_h - 1)} - 1} \quad (16)$$

Given that $\gamma_h > \sigma_h - 1$, a proportional decrease of fixed cost f_{ij}^{h} increases the total value of exports, X_{ij}^{h}, less than a proportional increase of the consumers' preference parameter α_{ij}^{h}. Because both α_{ij}^{h} and some elements of fixed cost f_{ij}^{h} are unobservable, we are not able to identify if the sequential and spatial exporting effect discovered in our empirical analysis is due to spatial correlated structure of fixed cost f_{ij}^{h} or to the spatial correlated preferences (the α_{ij}^{h} parameter). One of the two variables, or probably both, are the source of the defined *extended gravity* (Morales et al., 2017). In appendix A, we test if the impact of the ratio $\alpha_{ij}^{h} \left(\frac{\alpha_{ij}^{h}}{f_{ij}^{h}}\right)^{\frac{\gamma_h}{(\sigma_h - 1)} - 1}$ of Equation 16 can be negligible. We conclude that by omitting preference and cost parameters, it is impossible to correctly explain the structure of trade, especially for zero trade and large values of X_{ij}^{h}.

3.1 Introducing spatially correlated preferences

Several scholars in marketing and empirical economics have found that spatial interdependence among individual consumers plays a critical role in consumers' preferences. Frenkel et al. (2002) apply a cross-border structure to identify the segmentation of international markets. They introduce a spatial association and a spatial contiguity model in the segmentation literature, departing from the classical spatial independence or countries-as-segment assumption. Using survey data on a store image measurement instrument, they find superior performance of the spatial contiguity and association models and a relative preference of the spatial independence model over the countries-as-segment model, showing that preferences are correlated and cut across national borders.

Yang and Allenby (2003), using a Bayesian spatial autoregressive discrete choice model, show how preferences for Japanese-made cars are related to geographically and demographically defined networks. The authors display as the autoregressive specification reflects patterns of heterogeneity where

[19] Total export are equal to: $X_{ij}^{h} = w_i L_i * \Pr(\varphi > \tilde{\varphi}_h) * E(x_{ij}^{h} | \varphi > \tilde{\varphi}_h)$

influence propagates within and across networks. They demonstrate that preferences and choice behavior are influenced by consumer's own tastes and the tastes of others. People who identify themselves with a particular group often adopt the preferences of the group, resulting in choices that are interdependent. Examples include the preference for particular brands (e.g. Abercrombie and Fitch) or even entire product categories, as minivans. Yang and Allenby (2003), computing a measure of physical proximity[20] and a demographic neighbors variable[21], prove that geographically defined networks are more important to explain individual consumer behaviors than demographic networks.

Rossi et al. (2005), studying rating data, find that the latent preference variable is subject to respondent-specific location and scale shifts. Their latent rating provides superior information on purchases than the traditional centering method. Bell and Song (2007) show that consumers' decisions to adopt a new Internet service is affected by interactions with other consumers who live in the same postal code area, confirming the previous findings that consumers' preferences are spatially dependent.

Using Google trend data at national and local levels, we provide some intuitive graphical evidences about spatial correlation of preferences. Google trend data reports the relative frequency of search of a random sample of users in a selected area and time range. Data are scaled on a range from 0 to 100 based on a topic's proportion to all searches on all selected topics[22]. Topic's content is based on searched words, or set of words, that are then categorized by Google.

In Figure 1 and Figure 2, we observe the relative frequency of search for a subset of leading motor vehicle manufacturers. On a national basis, consumers in closer countries seem to search more for the same car producer, following a spatial pattern. This phenomenon is more evident when we move to a local level (Panels b of Figures 1 and 2). Switzerland is a multicultural country in the middle of Europe strongly influenced by habits, languages, culture and migration from their neighbours where preferences seem to be shaped along the border, as shown in panel b of Figure 1. Consumers in regions close to France search more for the French automaker Peugeot, consumers close to Italy for the Italian automaker FIAT and consumers close to Germany and Austria for the German automaker OPEL. Zurich and some

---

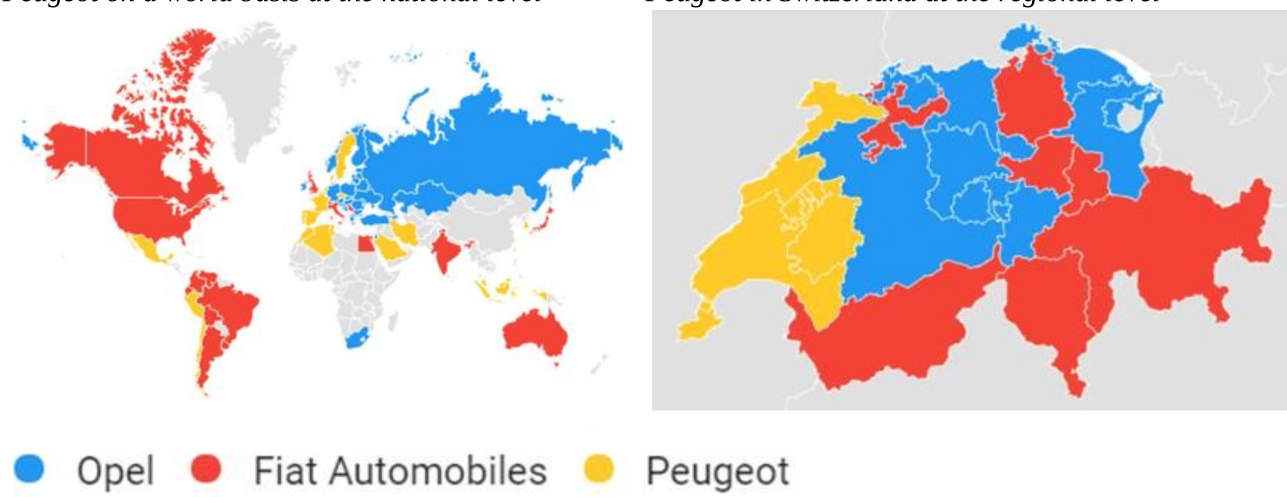
[20] Measured in terms of geographic distance among individuals' places of residence.
[21] Defined by membership in the same cluster with similar individuals.
[22] Google popularity of i in j = 100* $\frac{(\text{Google searches including the world or topic i/total Google searches})_j}{(\text{Google searches including the world or topic i/total Google searches})_{max}}$, where i is the searched word or topic and j is the local area.

others cantons (local entity at NUTS 2 levels) exhibit a more differentiated pattern of relative search, with a small preference for FIAT.

Figure 1 – Popularity (google relative trend data) for Opel, Fiat and Peugeot at a national and regional (Switzerland) level, for the year 2014-2015.

a) Relative frequency of search for Opel, Fiat and Peugeot on a world basis at the national level

b) Relative frequency of search for Opel, Fiat and Peugeot in Switzerland at the regional level

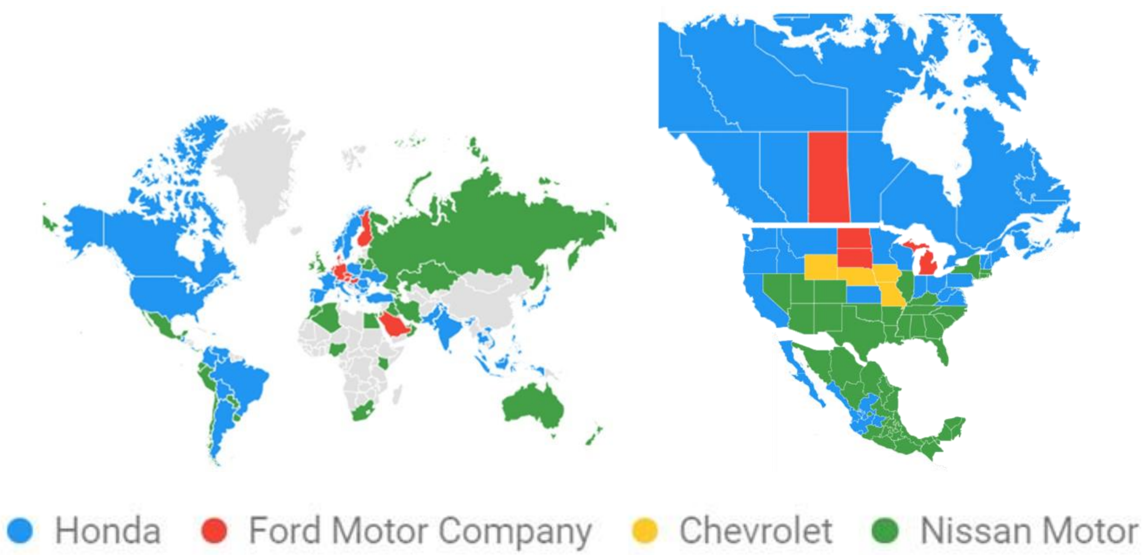


Note: Colour of regions or nations indicates the relative most searched carmaker in the years 2014-2015.

Figure 2 – Popularity (google relative trend data) for Honda, Ford, Chevrolet and Nissan at a national and regional (North America) level, for the year 2014-2015

a) Relative frequency of search for Honda, Ford, Chevrolet and Nissan on a world basis at the national level

b) Relative frequency of search for Honda, Ford, Chevrolet and Nissan in North America at the regional level



Note: Colour of regions or nations indicates the relative most searched carmaker in the years 2014-2015.

Similar considerations emerge from Figure 2, panel b. Relative searches popularity in North America clearly cross national borders and seem to be spatially correlated. Local consumers are more likely to look for the same car producer searched by their regional neighbours.

With the Moran's I test, we verify the presence of spatial correlation in the national data of Figure 1. As shown in Table 3, modelling preferences as spatially correlated seem to be justified, given that the test reject at 95% the null hypothesis of zero spatial autocorrelation. Z-score greater than zero points us that high or low values are more spatially clustered than expected[23].

Table 3 - Moran's I measure of spatial autocorrelation of the Google popularity index for FIAT, Peugeot and Opel carmakers in 2014, at a national level.

Variables	*I*	*E(I)*	*sd(I)*	*z*	*p-value*
GPI_{FIAT}	*0.050*	*-0.019*	*0.032*	*2.163*	*0.015*
GPI_{OPEL}	*0.339*	*-0.019*	*0.033*	*10.968*	*0.000*
$GPI_{PEUGEOT}$	*0.132*	*-0.019*	*0.032*	*4.642*	*0.000*

Note: This table shows the result of the Moran's I test of spatial autocorrelation. Based on these results, we can reject the null hypothesis that there is zero spatial autocorrelation in the variables at alpha = .05. I is the Moran I statistic, E(I) the expected value of the statistic under the null hypothesis of global spatial independence, sd(I) the standard deviation of the statistic, z the z-value of the statistic and p-value the corresponding 1-tail value.

Including spatial demand in a theoretical model is not straightforward. From a microeconomic point of view, we define the preference parameter α of consumer *e* for good *h* produced by *i* as $\alpha_{ie}^{h} = z_{ie}\phi + \sum_{f\neq e}^{E} \frac{1}{f(d_{ef})} \beta_{ef} \alpha_{if}^{h}$ where $z_{ie} = (g_i, r_e)$, is a vector of parameters including features of good *h* produced in country *i* (g_i) and characteristics of the consumer *e* (r_e), *f* is a consumers different from *e*, $\frac{1}{f(d_{ef})}$ is a function of distance expressing the probability that the two consumers get in touch (it can be a physical distance or a network linked matrix), β_{ef} is the individual weight that *e* gives to the preferences of *f*, capturing the imitation behavior of *e* and the social proximity between *e* and *f*[24].

---

[23] These seem compatible with the findings of Appendix A. Omitting the unobserved preference and fixed cost parameters when trade is predicted by a gravity equation produces large bias especially for low and high value of the distribution.

[24] A more formal micro approach is used by Yang and Allenby (2003), where the binary choice of a good captures the potential social dependency of preferences among consumers. The latent preference of *i* for good 2 over 1, defined as z_i is captured by $z_i = x_i'\beta + \varepsilon_i + \theta_i$, with $\theta = \rho W\theta + u$, $\varepsilon \sim N(0, I), u \sim N(0, \sigma^2 I)$, where θ is a vector of autoregressive parameter with ρW capturing the interdependence of preferences across consumers, and W is a matrix of finite mixture of coefficient, $W = \sum_{k=1}^{K} \phi_k W_k$, with $\sum_{k=1}^{K} \phi_k = 1$ where k are factors capturing the theoretical proximity of consumer (k_1 could be the physical distance, k_2 the wealth distance, k_3 ethnicity and so on).

National average preferences are therefore equal to the mean of α_{ie}^h over all consumers *e* belonging to country *j*. Formally, we simplify the model defining the average preference of consumers in country *j* for the good *h* with country-of-origin *i*, α_{ij}^h, as

$$\alpha_{ij}^h = z_{ij}^h \phi + \sum_{k \neq j}^{K} \frac{\beta_{jk}^h}{d_{jk}^h} \alpha_{ik}^h \tag{17}$$

where z_{ij}^h is a vector including the average characteristics of consumers in *j* and the features of product *h* produced in *i*, β_{jk} is the average influence parameter of consumer in country *k* over consumers in country *j* and d_{jk}^h is a function of the distance between consumer in country *j* and *k* (e.g. population weighted distance or a border binary variable) controlling for the probability that consumers in countries *j* and *k* are in contact or they observe each other[25].

Considering the spatial dependence of the preference parameter α_{ij}^h, we can therefore rewrite Eq. 16 as:

$$X_{ij}^h = \mu_h \frac{Y_i\, Y_j}{Y} \left(\frac{w_i \tau_{ij}^h}{\theta_j^h}\right)^{-\gamma_h} \left(z_{ij}^h \phi + \sum_{k \neq j}^{K} \frac{\beta_{jk}^h}{d_{jk}^h} \alpha_{ik}^h\right)^{\frac{\gamma_h}{(\sigma_h - 1)}} \left(f_{ij}^h\right)^{1-\frac{\gamma_h}{(\sigma_h - 1)}} \tag{18}$$

and plugging α_{ik}^h derived from Eq. 16[26] as a function of the observable parameter X_{ik}^hin Equation 18:

$$X_{ij}^h = \mu_h \frac{Y_i\, Y_j}{Y} \left(\frac{w_i \tau_{ij}^h}{\theta_j^h}\right)^{-\gamma_h} \left(z_{ij}^h \phi + \left(\frac{Y {w_i}^{\gamma}}{\mu_h Y_i}\right)^{\frac{\sigma_h - 1}{\gamma}} \sum_{k \neq j}^{K} \frac{\beta_{jk}^h}{d_{jk}^h} \left[\left(\frac{X_{ik}^h}{Y_k}\right)^{\frac{1}{\gamma_h}} \left(\frac{\tau_{ik}^h}{\theta_k^h}\right) \left(f_{ik}^h\right)^{\left(\frac{1}{\gamma_h} - \frac{1}{\sigma_h - 1}\right)}\right]^{\sigma_h - 1}\right)^{\frac{\gamma_h}{(\sigma_h - 1)}} \left(f_{ij}^h\right)^{1-\frac{\gamma_h}{(\sigma_h - 1)}} \tag{19}$$

[25] Consequently, large area countries with sparse population and low accessibility will exhibits less homogenous preferences.

[26] Using Eq. 16, $\alpha_{ij}^h = \left(\frac{Y X_{ij}^h}{\mu_h Y_i\, Y_j}\right)^{\frac{\sigma_h - 1}{\gamma_h}} \left(\frac{w_i \tau_{ij}^h}{\theta_j^h}\right)^{\sigma_h - 1} \left(f_{ij}^h\right)^{\frac{\sigma_h - 1}{\gamma_h} - 1}$

Ceteris paribus, exports from country *i* to countries different from *j* are positively correlated with export from *i* to *j*, if the influence parameter (β_{jk}) is greater than 0. This correlation is higher for countries close to country *j* (because of d_{jl}^{h}), as we have shown in our empirical analysis (reduced from).

This *spatially demand gravity equation*, which includes in the traditional gravity framework a spatially correlated preference parameter, explains some of the empirical findings of previous scholars, such as Chaney (2014), Morales et al. (2017) or Blum and Goldfarb (2006) [27] and provides a useful framework to explore the migration-trade link effect[28]. Indeed, plugging Equation 17 into the equation of firm's exports (Equation 15) we can write

$$x_{ij}(\varphi) = \begin{cases} \Phi_2 \left(\frac{Y_j}{Y}\right)^{\frac{\sigma_h - 1}{\gamma_h}} \left(z_{ij}^h \phi + \sum_{k \neq j}^{K} \frac{\beta_{jk}^h}{d_{jk}^h} \alpha_{ik}^h \right) \left(\frac{w_i \tau_{ij}^h}{\theta_j^h} \right)^{1-\sigma_h} \varphi^{\sigma_h - 1} & if\ \varphi \geq \tilde{\varphi}_{ij}^h \\ 0 & if\ \varphi < \tilde{\varphi}_{ij}^h \end{cases} \tag{20}$$

with threshold

$$\tilde{\varphi}_{ij}^h = \Phi_3 \left(\frac{Y}{Y_j}\right)^{\frac{1}{\gamma_h}} \frac{w_i \tau_{ij}^h}{\theta_j^h} \left(\frac{f_{ij}^h}{z_{ij}^h \phi + \sum_{k \neq j}^{K} \frac{\beta_{jk}^h}{d_{jk}^h} \alpha_{ik}^h} \right)^{1/(\sigma_h - 1)} \tag{21}$$

[27] Blum and Goldfarb (2006) find distance impact lasts in the case of digital goods consumed over the Internet that have no trading costs. They show that the effect of distance only holds for taste-dependent digital products, such as music, games, and pornography, while disappears for non-taste-dependent products.

[28] Two main channels have been described in the literature to explain how immigrants can enhance trade: the information/search cost channel (fixed cost) and the transaction cost channel (variable cost). Migrants can serve as information providers and trade intermediaries because they have a deep knowledge of their home country's opportunities and potential markets, access to distribution channels, contacts and familiarity to local customs, law and business practices. In our framework, we overtake the fixed or variable cost migration effect, introducing a realistic impact on national consumers' preferences. We can therefore overhaul the traditional debate on the migration trade link defining a theoretical framework where immigrant can influence both the fixed cost to export (reducing information cost) and the preference parameter of consumers in the importing country, without resort to shape variable trade cost as a function of workers or consumers' characteristics. The idea that migrants reduce the variable cost of trade has been proposed in order to explain why immigration increaes the intensive margin of trade.

Deriving α_{ik}^h in Equation 20 as a function of the observable parameter, X_{ik}^h we have:

$x_{ij}(\varphi)$ =

$$= \begin{cases} \Phi_2 \left(\frac{Y_j}{Y}\right)^{\frac{\sigma_h - 1}{\gamma_h}} \left(z_{ij}^h \phi + \left(\frac{Y w_i^{\gamma}}{\mu_h Y_i}\right)^{\frac{\sigma_h - 1}{\gamma}} \sum_{k \neq j}^{K} \frac{\beta_{jk}^h}{d_{jk}^h} \left[\left(\frac{X_{ik}^h}{Y_k}\right)^{\frac{1}{\gamma_h}} \left(\frac{\tau_{ik}^h}{\theta_k^h}\right) \left(f_{ik}^h\right)^{\left(\frac{1}{\gamma_h} - \frac{1}{\sigma_h - 1}\right)} \right]^{\sigma_h - 1} \right) \left(\frac{w_i \tau_{ij}^h}{\theta_j^h}\right)^{1-\sigma_h} \varphi^{\sigma_h - 1} & if\ \varphi \geq \tilde{\varphi}_{ij}^h \\ 0 & if\ \varphi < \tilde{\varphi}_{ij}^h \end{cases} \quad (22)$$

with threshold

$$\tilde{\varphi}_{ij}^h = \Phi_3 \left(\frac{Y}{Y_j}\right)^{\frac{1}{\gamma_h}} \frac{w_i \tau_{ij}^h}{\theta_j^h} \left(\frac{f_{ij}^h}{z_{ij}^h \phi + \left(\frac{Y w_i^{\gamma}}{\mu_h Y_i}\right)^{\frac{\sigma_h - 1}{\gamma}} \sum_{k \neq j}^{K} \frac{\beta_{jk}^h}{d_{jk}^h} \left[\left(\frac{X_{ik}^h}{Y_k}\right)^{\frac{1}{\gamma_h}} \left(\frac{\tau_{ik}^h}{\theta_k^h}\right) \left(f_{ik}^h\right)^{\left(\frac{1}{\gamma_h} - \frac{1}{\sigma_h - 1}\right)} \right]^{\sigma_h - 1}} \right)^{1/(\sigma_h - 1)} \quad (23)$$

$\tilde{\varphi}_{ij}^h$ is lower for countries *j* that are closer in term of distance (lower d_{jk}^h) or in term of similarity (higher influence parameter β_{jk}) to countries where *i* already exports. An increase of β_{jk}^h or $z_{ij}^h \phi$[29], due for instance to migration or to increasing proximity with other consumers in *k*, reduces the minimum level of productivity to access market *j* and boosts the value of trade for each firm. We will therefore observe an increase in the extensive margin of trade, that is defined by $pr(\varphi > \tilde{\varphi}_{ij}^h)$ and in the intensive margin of trade (x_{ij}). So far, we are able to explain the trade-migration effect on the intensive margin of trade without imposing a particular structure to variable trade costs.

With spatial correlated preferences, consumers prefer to buy goods used by their geographical or social neighbours and this effect is independent from logistic or supply chain costs included in the fixed cost. This effect is stronger and explicit along the border, because connections and social interactions decrease rapidly with distance.

[29] $z_{ij}^h \phi$ is the average idiosyncratic preference parameter of consumers in country *j* for product *h* made by *i*, that is influenced by the features of customers in country *i*.

As found at a micro level by Chaney (2014) and at a macro level in our empirical analysis, the probability to observe trade flows between *i* and *j* increases with the total exports of *i* to countries *k* different from *j* ($\sum_{k\neq j}^{K} X_{ik}^{h}$) and decreases with the distance of country *j* from these other countries *k* (d_{jk}^{h}). According to the results of our reduced form, this effect is stronger for countries sharing some characteristics (the extended gravity variable), because of different β_{jk}^{h}.

4 Estimation of the correlation parameter

4.1 Identification strategy

In order to identify the preference parameter in the preference-fixed cost ratio, we need to cancel out the fixed cost variable. Exports from countries *i* and *k* to country *j*, for good *h*, are equals to $X_{ij}^{h} = \mu_h \frac{Y_i Y_j}{Y}\left(\frac{w_i \tau_{ij}^{h}}{\theta_j^{h}}\right)^{-\gamma_h} (\alpha_{ij}^{h})^{\frac{\gamma_h}{(\sigma_h-1)}} (f_{ij}^{h})^{1-\frac{\gamma_h}{(\sigma_h-1)}}$ and $X_{kj}^{h} = \mu_h \frac{Y_k Y_j}{Y}\left(\frac{w_k \tau_{kj}^{h}}{\theta_j^{h}}\right)^{-\gamma_h} (\alpha_{kj}^{h})^{\frac{\gamma_h}{(\sigma_h-1)}} (f_{kj}^{h})^{1-\frac{\gamma_h}{(\sigma_h-1)}}$

Taking the ratio of X_{ij}^{h} over X_{kj}^{h} we can write:

$$\frac{X_{ij}^{h}}{X_{kj}^{h}} = \frac{Y_i}{Y_k}\left(\frac{w_i \tau_{ij}^{h}}{w_k \tau_{kj}^{h}}\right)^{-\gamma_h} \left(\frac{f_{ij}^{h}}{f_{kj}^{h}}\right)^{1-\frac{\gamma_h}{(\sigma_h-1)}} \left(\frac{\alpha_{ij}^{h}}{\alpha_{kj}^{h}}\right)^{\frac{\gamma_h}{(\sigma_h-1)}} \tag{24}$$

Our identification strategy relies on the fact that if *i*, *k* and *j* belong to the same custom union or common market *S*, $\frac{f_{ij}^{h}}{f_{kj}^{h}} = 1$. Fixed costs to export of the two countries, f_{ij}^{h} and f_{kj}^{h}, must be equals, given that country *j* has to guarantee the same importing conditions for products of countries *i* and *k*. Therefore, firms in countries *k* and *i* face the same fixed costs to export to market *j*. This assumption holds as long as networking, marketing, logistic, procedural fixed costs and sunk cost are the same for the exporting firms producing in countries *i* and *k*. This is a reasonable condition until we consider countries close to each other, in custom or market union, where firms shall have access to the same set of market information for the same price.

In a custom union *S*, where *i*, *j* and $k \in S$, the ratio between the exports of countries *i* and *k* to country *j* can be written as:

$$\frac{X_{ij}^{h}}{X_{kj}^{h}} = \frac{Y_i}{Y_k}\left(\frac{w_i \tau_{ij}^{h}}{w_k \tau_{kj}^{h}}\right)^{-\gamma_h} \left(\frac{\alpha_{ij}^{h}}{\alpha_{kj}^{h}}\right)^{\frac{\gamma_h}{(\sigma_h-1)}} \tag{25}$$

Recalling that

$$\alpha_{ij}^h = z_{ij}^h \phi + \sum_{l \neq j}^L \frac{\beta_{jl}^h}{d_{jl}^h} \alpha_{il}^h \text{ and } \alpha_{kj}^h = z_{kj}^h \phi + \sum_{l \neq j}^L \frac{\beta_{jl}^h}{d_{jl}^h} \alpha_{kl}^h \tag{26}$$

and, using Equation 25 to define α_{il}^h and α_{kl}^h as a function of the ratio between the observable exports from the same country *s* to country *l*, X_{sl}^h, over the export from country *i* to *l*, X_{il}^h, and *k* to *l*, X_{kl}^h (with $i,k,l,s \in S$)[30], we can write

$$\alpha_{il}^h = \left[\frac{X_{il}^h}{X_{sl}^h}\frac{Y_s}{Y_i}\left(\frac{w_s \tau_{sl}^h}{w_i \tau_{il}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h - 1}{\gamma_h}} \alpha_{sl}^h \text{ and } \alpha_{kl}^h = \left[\frac{X_{kl}^h}{X_{sl}^h}\frac{Y_s}{Y_k}\left(\frac{w_s \tau_{sl}^h}{w_k \tau_{kl}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h - 1}{\gamma_h}} \alpha_{sl}^h \tag{27}$$

Plugging Equation 27 into Equations 26 and substituting α_{ij}^h and α_{kj}^hin Equation 25 we have:

$$\frac{X_{ij}^h}{X_{kj}^h} = \frac{Y_i}{Y_k}\left(\frac{w_i \tau_{ij}^h}{w_k \tau_{kj}^h}\right)^{-\gamma_h} \left(\frac{z_{ij}^h \phi + \left(\frac{Y_s}{Y_i}\left(\frac{w_s}{w_i}\right)^{-\gamma_h}\right)^{\frac{\sigma_h - 1}{\gamma_h}} \sum_{l \neq j}^L \frac{\beta_{jl}^h}{d_{jl}^h}\left[\frac{X_{il}^h}{X_{sl}^h}\left(\frac{\tau_{sl}^h}{\tau_{il}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h - 1}{\gamma_h}} \alpha_{sl}^h}{z_{kj}^h \phi + \left(\frac{Y_s}{Y_k}\left(\frac{w_s}{w_k}\right)^{-\gamma_h}\right)^{\frac{\sigma_h - 1}{\gamma_h}} \sum_{l \neq j}^L \frac{\beta_{jl}^h}{d_{jl}^h}\left[\frac{X_{kl}^h}{X_{sl}^h}\left(\frac{\tau_{sl}^h}{\tau_{kl}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h - 1}{\gamma_h}} \alpha_{sl}^h} \right)^{\frac{\gamma_h}{(\sigma_h - 1)}} \tag{28}$$

It is straightforward to note that for $\beta_{jl}^h > 0$, $\frac{dX_{ij}^h}{dX_{il}^h}$ is always greater than 0. Because of spatial correlated demand, exports from country *i* to countries different from *j* are positively correlated with imports of country *j* from country *i*. This effect is independent from the supply effects of *i* and *k* and from variations in the fixed costs f_{ij}^h and f_{kj}^h, as we are controlling for wage cost and considering *i*, *k*, *l* and *s* belonging to the same custom union *S*[31].

If $\beta_{jl}^h > 0$, the value of the ratio $\frac{X_{ij}^h}{X_{kj}^h}$ is positively correlated with the ratio between $\sum_{l \neq j}^L \frac{X_{il}^h}{X_{sl}^h}$ over $\sum_{l \neq j}^L \frac{X_{kl}^h}{X_{sl}^h}$, holding constant all the other parameters.

---

[30] This approach holds even in absence of custom union, until countries in the group *S* have to guarantee the same importing conditions at each member of the group.

In order to identify the spatial correlation parameter, we rewrite Equation 28 introducing the time dimension. Recalling that the idiosyncratic preference parameter $z_{ij}^h\phi$ is a vector of features of good *h* produced by country *i* and characteristics of consumers in *j*, we can decompose z_{ij}^h in the vector component r_{jt}, capturing all the time varying characteristics of consumers in *j*, and g_i^h controlling for the time invariant features of good *h* produced by *i*. We set the idiosyncratic preference parameter $z_{ij}^h\phi$ and $z_{kj}^h\phi$ equal to $g_i^h\phi_j^h + r_{jt}\phi_0^h$ and $g_k^h\phi_j^h + r_{jt}\phi_0^h$ respectively. We can express these terms as fixed effects, with ψ_{ij}^h and ψ_{kj}^h capturing the time-invariant preferences of consumers in *j* for the goods produced by *i* and *k*, and ψ_{jt}^h controlling for the time varying preferences of *j* for good *h*. Considering $\beta_{jl}^h = \beta^h \; \forall j, l \in S$ we have:

$$\frac{X_{ijt}^h}{X_{kjt}^h} = \frac{Y_{it}}{Y_{kt}}\left(\frac{w_{it}\tau_{ijt}^h}{w_{kt}\tau_{kjt}^h}\right)^{-\gamma_h}\left(\frac{\psi_{ij}^h + \psi_{jt}^h + \left(\frac{Y_{st}}{Y_{it}}\left(\frac{w_{st}}{w_{it}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}}\beta^h\sum_{l\neq j}^{L}\frac{\alpha_{slt}^h}{d_{jl}^h}\left[\frac{X_{ilt}^h}{X_{slt}^h}\left(\frac{\tau_{slt}^h}{\tau_{ilt}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h-1}{\gamma_h}}}{\psi_{kj}^h + \psi_{jt}^h + \left(\frac{Y_{st}}{Y_{kt}}\left(\frac{w_{st}}{w_{kt}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}}\beta^h\sum_{l\neq j}^{L}\frac{\alpha_{slt}^h}{d_{jl}^h}\left[\frac{X_{klt}^h}{X_{slt}^h}\left(\frac{\tau_{slt}^h}{\tau_{klt}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h-1}{\gamma_h}}}\right)^{\frac{\gamma_h}{\sigma_h-1}} \tag{29}$$

The idiosyncratic preferences are therefore a function of bilateral static preferences of consumers in *j* for product *h* produced by *i* or *j* and a time varying preference of consumers in *j* for the product *h*. The assumption of static idiosyncratic bilateral preference is needed to better identify our parameters and holds considering not too long length of time. In Appendix B some concerns about possible endogeneity is examined in greater depth.

4.2 Model estimation

In order to estimate constrained and nested parameters and to avoid incidental parameter problem, we employ MCMC methods. The number of parameters to estimate and the number of observations are a function of timespan and number of countries in the custom union we consider. The higher the number of countries in the custom union, the better the probabilities to correctly identify the parameter β^h of our equation. To evaluate the performance of our MCMC estimator, we run some simulations considering ten years of trade in a custom union with ten members. Results are in Appendix C.

To estimate the parameters of our model, we resort to some simplifying assumptions. We log linearized Equation 29 and we assume the log of exports' ratio (our dependent variable) being distributed as a normal with mean μ and independent error ε_{ijkt}^{h} with variance δ_h. Formally:

$$\ln\left(\frac{X_{ijt}^h}{X_{kjt}^h}\right) \sim N(\mu, \delta_h)$$

$$\mu = ln\left(\frac{Y_{it}}{Y_{kt}}\right) - \gamma_h ln\left(\frac{w_{it}\tau_{ijt}^h}{w_{kt}\tau_{kjt}^h}\right) +$$

$$+\left(\frac{\gamma_h}{\sigma_h - 1}\right) ln\left(\frac{\psi_{ij}^h + \psi_{jt}^h + \left(\frac{Y_{st}}{Y_{it}}\left(\frac{w_{st}}{w_{it}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}} \beta^h \sum_{l\neq j}^{L} \frac{\alpha_{slt}^h}{d_{jl}^h}\left[\frac{X_{ilt}^h}{X_{slt}^h}\left(\frac{\tau_{slt}^h}{\tau_{ilt}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h-1}{\gamma_h}}}{\psi_{kj}^h + \psi_{jt}^h + \left(\frac{Y_{st}}{Y_{kt}}\left(\frac{w_{st}}{w_{kt}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}} \beta^h \sum_{l\neq j}^{L} \frac{\alpha_{slt}^h}{d_{jl}^h}\left[\frac{X_{klt}^h}{X_{slt}^h}\left(\frac{\tau_{slt}^h}{\tau_{klt}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h-1}{\gamma_h}}}\right) \tag{30}$$

with $\gamma_h > \sigma_h > 1$ and $\psi_{ij}^h, \psi_{jt}^h, \alpha_{slt}^h > 0$, as specified in the theoretical setting[32].

Log linearization is a standard approach in international trade literature and allows us to efficiently deal with the numeric optimization of the sampler, reducing the scale. To implement the estimator, we specify prior density functions for the unknown parameters; combining these with the likelihood function of our equation and dividing them by the marginal distribution of the data we obtain the posterior distribution of our parameters[33]. We then sample from this distribution using a No-U-Turn sampler algorithm (Hoffman et al., 2014).

In order to avoid discontinuity or infeasible value of μ in the estimation process, we constrain parameter β^h to be equal or greater than zero. All the priors are therefore modeled as exponential or function of exponential (Appendix D). Trade flows data comes from UN-Comtrade, for the period 2002-2012, and are collected at 2 digit HS reported level (97 sectors) with annual frequency. For our identification purpose, we chose to consider the biggest available set of nations included in a custom union: the European countries. To avoid issued related to different currencies, which could bias our estimates (because of potential different fixed costs to export linked to financial cost) we further restrict

[32] Because we are considering only positive trade, the constrain $\alpha_{ij}^h \geq 0$ becomes $\alpha_{ij}^h > 0$. Consequently ψ_{ij}^h, ψ_{jt}^h are set greater than zero because $\alpha_{ijt}^h = \psi_{ij}^h + \psi_{jt}^h$

[33] Formally $\frac{p(y|\theta)p(\theta)}{\int_\theta p(y|\theta)p(\theta)d\theta}$ where $y = ln\left(\frac{X_{ijt}^h}{X_{kjt}^h}\right)$, $p(y|\theta) = (2\pi\sigma^2)\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \mu)^2\right)$ and θ is our set of parameters.

the set to 11 countries belonging to the Euro area: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal and Spain, with Finland sets as the reference country s. These countries not only belong to the same custom union but are also members of a common market, strengthening our identification strategy[34].

We resort to the following sources for the variables of the model: wages data (hourly compensation costs in manufacturing) come from U.S. Bureau of Labor Statistics, International Labor Comparisons, August 2013; population weighted distances are from CEPII and national GDP data from Penn World Table. Sector variable trade costs (τ_{ilt}^{h}) at HS 2 digit level are computed from OECD ITIC database[35] (Miao and Fortanier, 2017).

Due to intensive computation, we proceed to estimate our parameters only for a subset of sectors. Below, we chose to report results for the HS 22 sector (Beverages, Spirits and Vinegar), assuming this sector produces some of the most differentiated goods with low spatial correlation of preferences. As shown in Table 4, β^{h} is positive and significantly different from zero.

[34] Choosing Finland as reference country, we try to reduce the correlation of parameter in the estimation process, reducing the time required for the MCMC procedure to explore efficiently the support of all the parameters. Considering that $\alpha_{slt}^{h} = \psi_{sl}^{h} + \psi_{lt}^{h} + \beta_{s}^{h} \sum_{i \neq l}^{K} \frac{1}{d_{li}} \left(\psi_{si}^{h} + \psi_{it}^{h} \right) + \beta_{s}^{h} \sum_{i \neq l}^{L} \frac{1}{d_{li}} \varepsilon_{sit}^{h} + \varepsilon_{slt}^{h}$ we might assume that consumers' preferences for the goods produced by Finnish firms are affected at a different level by the neighbours' preferences. If $\beta^{h} \neq \beta_{s}^{h}$, the lower correlation between the parameter can lead to shorter runtimes to achieve the true joint posterior distribution.
Anyway, the testing hypothesis $\beta^{h} \neq 0$ is always valid, because if $\beta^{h} \sum_{i \neq l}^{L} \frac{1}{d_{li}} \alpha_{slt}^{h} = 0 \implies \beta^{h} = 0$ given that α_{slt}^{h} is greater than 0 for each s ,l ,t, since we are considering only positive trade flows.

[35] http://oecdinsights.org/2016/11/02/statistical-insights-new-oecd-database-on-international-transport-and-insurance-costs/

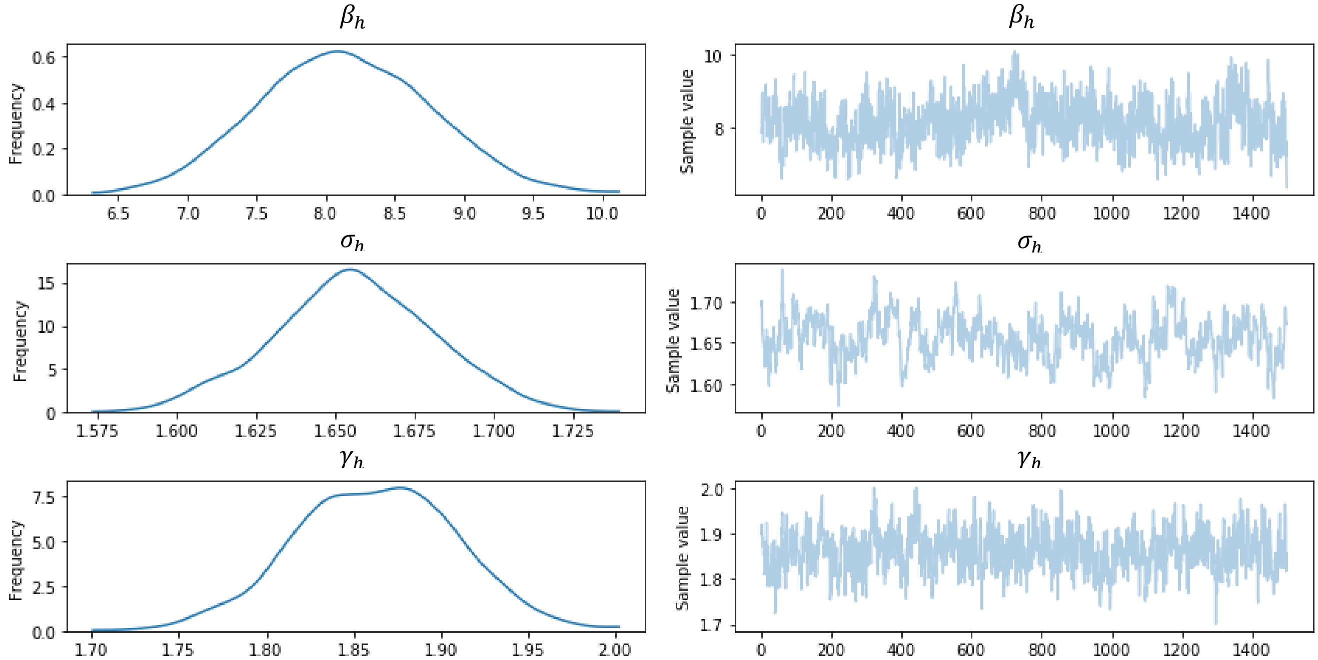
Table 4 – Parameters estimation of Equation 30 for the HS 22 Sector (Beverages, Spirits and Vinegar)

	Median	SD	MC Error	95% HPD Interval
β^h	37.87	3.694	0.120	[30.96, 45.33]
σ_h	1.80	0.064	0.003	[1.678, 1.924]
γ_h	1.851	0.065	0.002	[1.723, 1.974]

Note: This table shows the estimated values of the spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 29 for sector HS 22. Values are computed using Hamiltonian Monte Carlo method (No-U-Turn Sampler), with a target acceptance rate of 0.95 for the sampler.

Inspecting the following Figure 3, which displays the distributions of the posterior means and traces of parameters of Table 4, we observe estimates converging to their medians.

Figure 3 – Posterior means and traces of parameters β^h, σ_h and γ_h of Table 3, for the HS 22 Sector (Beverages, Spirits and Vinegar)



Note: In Figure 3, we observe the distribution of the estimated spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 30. Values are computed using Hamiltonian Monte Carlo method (No-U-Turn sampler), with a target acceptance rate of 0.95 for the sampler, discarding the first 1000 draws.

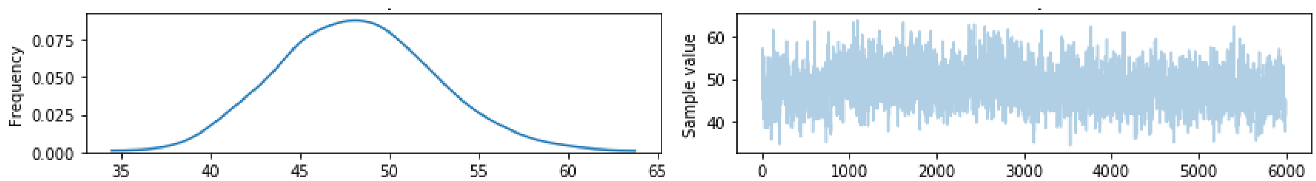
As shown in Figure 3, consumers' preferences are spatially correlated, at least for the HS 22 sector and for the consumers in the countries included in our analysis. However, the strength of this spatial correlation is not very high. As expected, beverage market is local and slightly influenced by neighbours' preferences, at least at a national level. Only very close consumers affect preferences and the effect rapidly diminishes increasing the distance. For Beverages, only consumers in small and close countries share consumption preferences.

However, other products provide stronger results. As shown in Figure 4, the spatial correlation parameter of preferences for Apparel, Chemical products and Vehicles is about 7 times bigger than for Beverages. Only Ceramic products provide an estimate that is comparable to the Beverages sector. These results suggest that some products are sold worldwide and others not, because preferences can be focused on a local or global scale, depending by the product we consider. According to our results, beer, vinegar or ceramic wares are perceived as local products while car or fashion brands as global goods. This should explain because companies in different sectors exhibit diverse marketing and development strategies. As example, corporations in the beer industry tend to expand their market shares buying local producers and maintaining the original brand names, without to advertise their operations (AB InBev and Heineken own more than 240 local beer brands) while luxury, sport fashion brands or car producers promote their products globally.

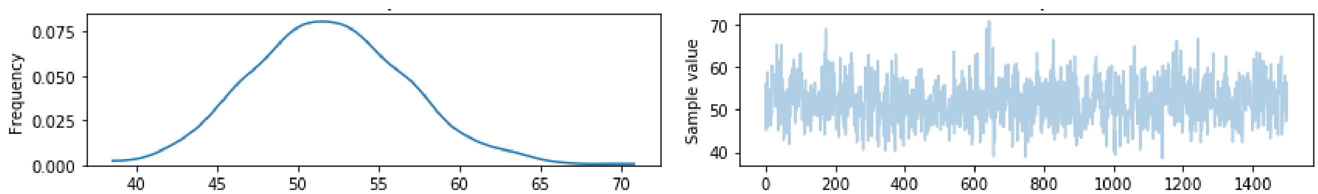
These outcomes are consistent with the findings of Davvetas and Diamantopoulos (2016). Measuring consumers' perceptions of global versus local brand superiority with two studies in developed and emerging markets, they prove that consumers rely on product category schemata to form perceptions. Their results imply that global or local brand preferences are largely formed at the product category level and consumers perceive global brands as superior to local brands in product categories with strong functional character and extensive symbolic capacity.

Figure 4 – Posterior means and traces of parameters β^h for the 61, 38, 69 and 87 HS sectors

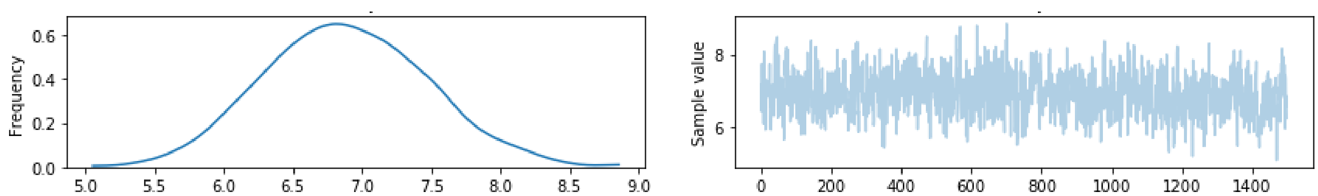
Sector 61 HS - Apparel and clothing accessories; knitted or crocheted.



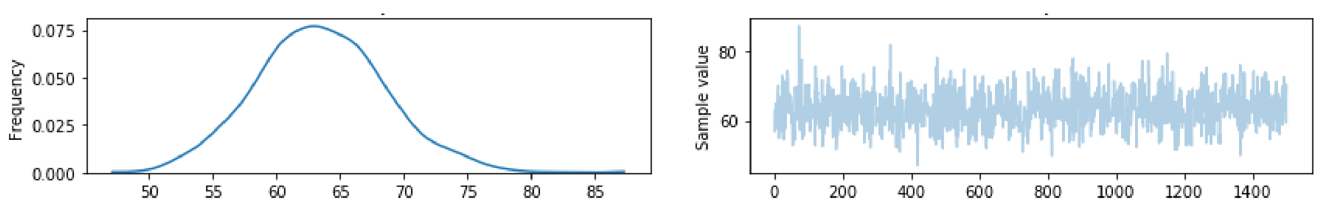
Sector 38 HS – Miscellaneous chemical products.



Sector 69 HS – Ceramic products.



Sector 87 HS – Vehicles; other than railway or tramway rolling stock.



Note: In Figure 4, we observe the distribution of the spatial correlation parameter (β^h) of equation 30 and its trace. Values are estimated using Hamiltonian Monte Carlo computation (No-U-Turn sampler), with a target acceptance rate of 0.95 (0.98 for HS61) for the sampler, discarding the first 1000 draws.

This simple way to model consumers' preferences is sufficient to prove that they are spatially correlated, confirming the assumption of correlated demand that is crucial for all the literature on search and learning models. In Appendix E we compute the average influence of neighbors using a formal standard matrix approach, with a row-normalized distance matrix and approximating the MCMC using automatic differentiation variational inference (ADVI, Kucukelbir et al., 2017). Results of simulations confirms neighbors' influence on preferences.

5 Conclusion

Starting from empirical evidences of spatially correlated exports at the extensive and intensive margin of trade, this paper suggests a development of the traditional gravity model of trade in order to take into account spatially correlated preferences.

We find reduced-form evidence of a positive correlation between bilateral trade and the spatial distribution of exports to other countries; both the probability to export to a target market and the value exported increase the more the exporting country sells its goods to the countries close to the target market. The probability to export is higher if the foreign consumers import similar products (belonging to the same SITC class) from countries already reached by the exporter.

Following several empirical findings in the marketing literature, we assume consumers in different countries share similar preferences and are influenced by consumption's decisions of their neighbours. Introducing spatially correlated preferences in the Chaney (2008) model of trade, we are able to explain our empirical findings. Modelling preferences as spatially dependents, we derive an extended aggregate equation of trade that can explain the "extended gravity" and "spatial exporters" effects discovered by previous scholars. Spatially correlated preferences are then confirmed in a structural estimation of our model for a subset of products and countries. We identify the spatial correlation parameter of consumers' preferences considering, in a custom union, the ratio of export to the same country from different countries, in order to control for observable and unobservable fixed costs to export.

To the best of our knowledge, we are the first to investigate explicitly the impact of preferences on international trade, modelling them as spatially dependents, and we are the first to stress the importance of the spatial structure of export from a demand perspective.

With this paper, we support search and learning models developed by previous authors, given that they assume demand as imperfectly correlated among countries. We even hope to encourage the integration of preference structuring into the international trading literature. Other directions are to explore the relationship between globalization and preferences or to assess the role of migration within this framework.

References

Albornoz, Facundo, Héctor F. Calvo Pardo, Gregory Corcos, and Emanuel Ornelas. 2012.“Sequential Exporting.” *Journal of International Economics 88 (1)*, 17–31.

Artopoulos, A., Friel, D., Hallak, J.C. 2011. Lifting the domestic veil: the challenges of exporting differentiated goods across the development divide. NBER Working Paper 16947

Blum, Bernardo S., Avi Goldfarb. 2006. “Does the internet defy the law of gravity?”. *Journal of International Economics*, Volume 70, Issue 2, Pages 384-405, ISSN 0022-1996, https://doi.org/10.1016/j.jinteco.2005.10.002.

Bell, David R., and Sangyoung Song. 2007. “Neighborhood effects and trial on the internet: Evidence from online grocery retailing.” *Quantitative Marketing and Economics Volume 5*, Issue 4, pp 361–400. DOI:https://doi.org/10.1007/s11129-007-9025-5

Bertoletti, Paolo, Federico Etro, and Ina Simonovska. 2018. "International Trade with Indirect Additivity." *American Economic Journal: Microeconomics*, 10 (2): 1-57.

Blei, David M. , Alp Kucukelbir, Jon D. McAuliffe. 2017. “Variational Inference: A Review for Statisticians”. Journal of the American Statistical Association, 112:518, 859-877, DOI: 10.1080/01621459.2017.1285773

Bradlow, Eric & Bronnenberg, Bart & Russell, Gary & Arora, Neeraj & Bell, David & Duvvuri, Sri Devi & Ter Hofstede, Frankel & Sismeiro, Catarina & Thomadsen, Raphael & Yang, Sha. 2005. “Spatial Models in Marketing.” *Marketing Letters*. 16. 267-278. 10.1007/s11002-005-5891-3.

Cavusgil, S. Tamer and Shaoming Zou. 1994. “Marketing Strategy-Performance Relationship: An Investigation of the Empirical Link in Export Market Ventures.” *Journal of Marketing, 58(Jan)*, 1-21.

Chaney, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review, 98*(4): 1707-21. DOI: 10.1257/aer.98.4.1707

Chaney, Thomas. 2014. "The Network Structure of International Trade." *American Economic Review*, 104(11): 3600-3634. DOI: 10.1257/aer.104.11.3600

Combes, P-P., Lafourcade M. and T. Mayer. 2005. “The Trade Creating Effects of Business and Social Networks: Evidence from France.” *Journal of International Economics*, 66(1): 1-29

Das, Sanghamitra, Mark J. Roberts, James R. Tybout, 2007. "Market Entry Costs, Producer Heterogeneity, and Export Dynamics." *Econometrica*, Econometric Society, vol. 75(3), pages 837-873, May.

Davvetas, Vasileios, and Adamantios Diamantopoulos, 2016. “How Product Category Shapes Preferences Toward Global and Local Brands: A Schema Theory Perspective.” *Journal of International Marketing*: December 2016, Vol. 24, No. 4, pp. 61-81.

Defever, F., Heid, B. & Larch, M. 2015. “Spatial exporters.” *Journal of International Economics*, 95(1), pp. 145-156. doi:10.1016/j.jinteco.2014.11.006

Dichter, E., 1962. The world customer. *Harvard Business Review*, 40 (4), 113-122.

Eaton, J., M. Eslava, M. Kugler and J. Tybout. 2008. “The Margins of Entry into Export Markets: Evidence from Colombia.” in E. Helpman, D. Marin and T. Verdier (eds.), *The Organization of Firms in a Global Economy,* Cambridge, MA: Harvard University Press, pp. 231-272.

Eaton, Jonathan, Marcela Eslava, David Jinkins, C. J. Krizan, and James R. Tybout. 2014. “A Search and Learning Model of Export Dynamics,”. 2015. Meeting Papers 1535, Society for Economic Dynamics.

Frenkel Ter Hofstede, Michel Wedel, and Jan-Benedict E.M. Steenkamp. 2002. “Identifying Spatial Segments in International Markets.” *Marketing Science 21*, 160-177.
DOI=http://dx.doi.org/10.1287/mksc. 21.2.167.154

Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo. 2005. “World Trade Flows: 1962-2000.” NBER Working Paper No. 11040.

Feenstra, Robert C., Philip Luck, Maurice Obstfeld, and Katheryn N. Russ. 2018. “In Search of the Armington Elasticity” The Review of Economics and Statistics 2018 100:1, 135-150

Garmendia, A., Llano-Verduras, C. and Requena-Silventre, F. 2012. “Network and the disappearance of the intranational home bias.” *Economic Letters 116*, 178-182.

Hoffman, Matthew D., and Andrew Gelman. 2014. “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.” *Journal of Machine Learning Research 15 (2014)* 1593-1623.

Jackson M.O., Rogers B.W. 2007. “Meeting strangers and friends of friends: how random are social networks?” *American Economic Review*, 97(3): 890 – 915

Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., and Blei, D. M. 2017. "Automatic differentiation variational inference." *Journal of the American Statistical Association Vol. 112*, Iss. 518, 2017. https://doi.org/10.1080/ 01621459.2017.1285773

Miao, G. and F. Fortanier. 2017. "Estimating Transport and Insurance Costs of International Trade." *OECD Statistics Working Papers*, No. 2017/04, OECD Publishing, Paris. http://dx.doi.org/10.1787/8267bb0f-en

Morales, E., Sheu, G., and Zahler, A. June 2017. "Extended Gravity". Revise and Resubmitted, *Review of Economic Studies*

Murphy, Kevin M, and Andrei Shleifer. 1997. "Quality and Trade." *Journal of Development Economics 53*, 1-15

Nguyen, Daniel X. 2012. "Demand uncertainty: Exporting delays and exporting failures." *Journal of International Economics*, Volume 86, Issue 2, Pages 336-344, ISSN 0022-1996, https://doi.org/10.1016/j.jinteco.2011.10.007.

Redding, S. and Venables, A. 2004. "Economic Geography and International Inequality". *Journal of International Economics*, 62, 53-82

Rossi, P., Allenby, G.M. and McCulloch, R. 2005. "Bayesian Statistics and Marketing". Wiley Series in Probability and Statistics, Chichester: Wiley.

Santos Silva, J.M.C. and Tenreyro, Silvana. 2006. "The Log of Gravity." *The Review of Economics and Statistics*, 88(4), pp. 641-658.

Stephens-Davidowitz, Seth. 2014. "The cost of racial animus on a black candidate: Evidence using Google search data." *Journal of Public Economics*, 118, issue C, p. 26-40. https://doi.org/10.1016/j.jpubeco.2014.04.010

Yang, Sha, Greg M. Allenby. 2003. "Modeling Interdependent Consumer Preferences." *Journal of Marketing Research*: August 2003, Vol. 40, No. 3, pp. 282-294.

Appendix

A Estimating the impact of unobservable fixed cost and consumers' preferences.

To estimate the impact of excluding the unobservable exogenous preference parameter α_{ij}^{h} and the unobservable component of the fixed cost f_{ij}^{h}, we compute the aggregate trade value X_{ij}^{h}, predicted by a classical gravity model.

Log-linearizing Equation 16, we obtain:

$$lnX_{ij}^{h} = ln(\mu_h) - \ln(Y) + \ln(Y_i) + \ln(Y_j) - \gamma_h \ln(\tau_{ij}^{h}) - \gamma_h \ln(w_i) + \gamma_h \ln(\theta_j^{h}) +$$

$$+\left(1 - \frac{\gamma_h}{(\sigma_h - 1)}\right) lnf_{ij}^{h} + \left(\frac{\gamma_h}{(\sigma_h - 1)}\right) ln\alpha_{ij}^{h} \qquad \text{(a.1)}$$

where fixed cost f_{ij}^{h} can be decomposed in observable fixed cost, $f_{ij}^{h}(obs)$, and unobservable fixed cost $f_{ij}^{h}(unobs)$ using the following specification:

$$f_{ij}^{h} = e^{f_{ij}^{h}(obs) + f_{ij}^{h}(unobs)} \qquad \text{(a.2)}$$

We assume observable fixed costs being a linear combination of a vector of parameters β and a set of covariates χ_{obs}, such as regional trade agreement (*rta*), spatial contiguity (*contig*), common language (*comlang*), common religion (*com_rel*, capturing cultural similarity), common currency (*com_cur*), colonial relationship (*col*), GATT/WTO membership (*gat_memb*)[36], that are proxies for fixed costs to export between countries *i* and *j*. In the following equation, we omit the coefficient parameters β for simplification.

$$f_{ij}^{h}(obs) = (rta_{ij} + contig_{ij} + comlang_{ij} + col_{ij} + com_{rel_{ij}} + com_{cur_{ij}} + gat_{memb_i} +$$

$$+gat\_memb_j) \qquad \text{(a.3)}$$

For each sector and year[37], we estimate the predicted value ($\hat{X}_{ij}^{h} = X_{ij}^{h} - \varepsilon_{ij}^{h}$) of the following equation (as above, we omit the parameters' vector and subfix *h* and *t* for simplification) using a PPML model (Santos Silva and Tenreyro, 2007):

[36] These variables are widely used in literature as a set of control variables for fixed costs.
[37] We perform separate regression for each sector and year.

$$X_{ij} = cost + \ln(gdp_i) + \ln(gdp_j) - \ln(dist_{ij}) + \theta_i + \theta_j + rta_{ij} + contig_{ij} + com\_lang_{ij} + col_{ij} + com\_rel_{ij} + com\_cur_{ij} + gat\_memb_i + gat\_memb_j + \varepsilon_{ij}^h \quad \text{(a.4)}$$

where θ_i captures all the observable (such as the Gdp level) and unobservable features of *i* influencing the trade flows in sector *h* at time *t* while θ_j captures all the observable and unobservable characteristics of *j*[38] in sector *h* at time *t*.

The error term ε_{ij}^h therefore captures the unobservable bilateral component of fixed cost, $f_{ij}^h(unobs)$, the bilateral exogenous preference parameter α_{ij}^h and an error term $\eta_{ij}^h \sim N(0, \delta_\eta)$.

$$\varepsilon_{ij}^h = X_{ij}^h - \hat{X}_{ij}^h = \left(1 - \frac{\gamma_h}{(\sigma_h - 1)}\right)\left(\ln e^{f_{ij}^h(unobs)} - \ln e^{f_i^h(unobs)} - \ln e^{f_j^h(unobs)}\right) - \left(\frac{\gamma_h}{(\sigma_h - 1)}\right)\left(\ln \alpha_{ij}^h - \ln \alpha_j^h - \ln \alpha_i^h\right) + \eta_{ij}^h \quad \text{(a.5)}$$

Using country fixed effects to control for the unobservable features of countries *i* and *j* in sector *h* (given that we are estimating the model for each sector and year) we are already controlling for average preferences and average fixed cost to export and import for each country. The residual α_{ij}^h and $e^{f_{ij}^h}$ in ε_{ij}^h are therefore the specific bilateral deviation of the country's preferences and unobservable fixed costs to export.

Accordingly to the source of our data, we set $\hat{X}_{ij}^h$ equal to one if the predicted value of trade is greater than zero but lower than one[39], in order to make predicted and real trade flows comparable. A low difference between our theoretical prediction and the real value will point out as negligible the effect of the unobserved bilateral preferences and fixed costs to export. Otherwise, we should specify some functional form for α_{ij}^h and f_{ij}^h to explain the distribution of the difference. As a first check, we verify if the traditional gravity model correctly predicts if country *i* exports good *h* to country *j*. For each trading pair, we compute a binary variable equal to 1 if the predicted trade flow for the good *h* is greater than 1 and 0 otherwise.

Results in Table A.1, show that the equation predicts very well if country *i* exports the good *h* to country *j*. The accuracy of the prediction (true positive) is about 99% and the rate of false positive is about 1%, however several problems arise for the prediction of zero trade flows. Observing a false positive rate of

[38] As Redding and Venables (2004), we estimate a theoretical founded gravity equation with fixed effects specification.
[39] Trade flows for each good between countries are in nominal thousands of US dollars ($1,000) and reported only if greater than 1,000 USD.

about 52%, we express some concerns that the omitted variables $(\alpha_{ij}^h, f_{ij}^h)$ could play a substantial role in the selection of the exporting markets.

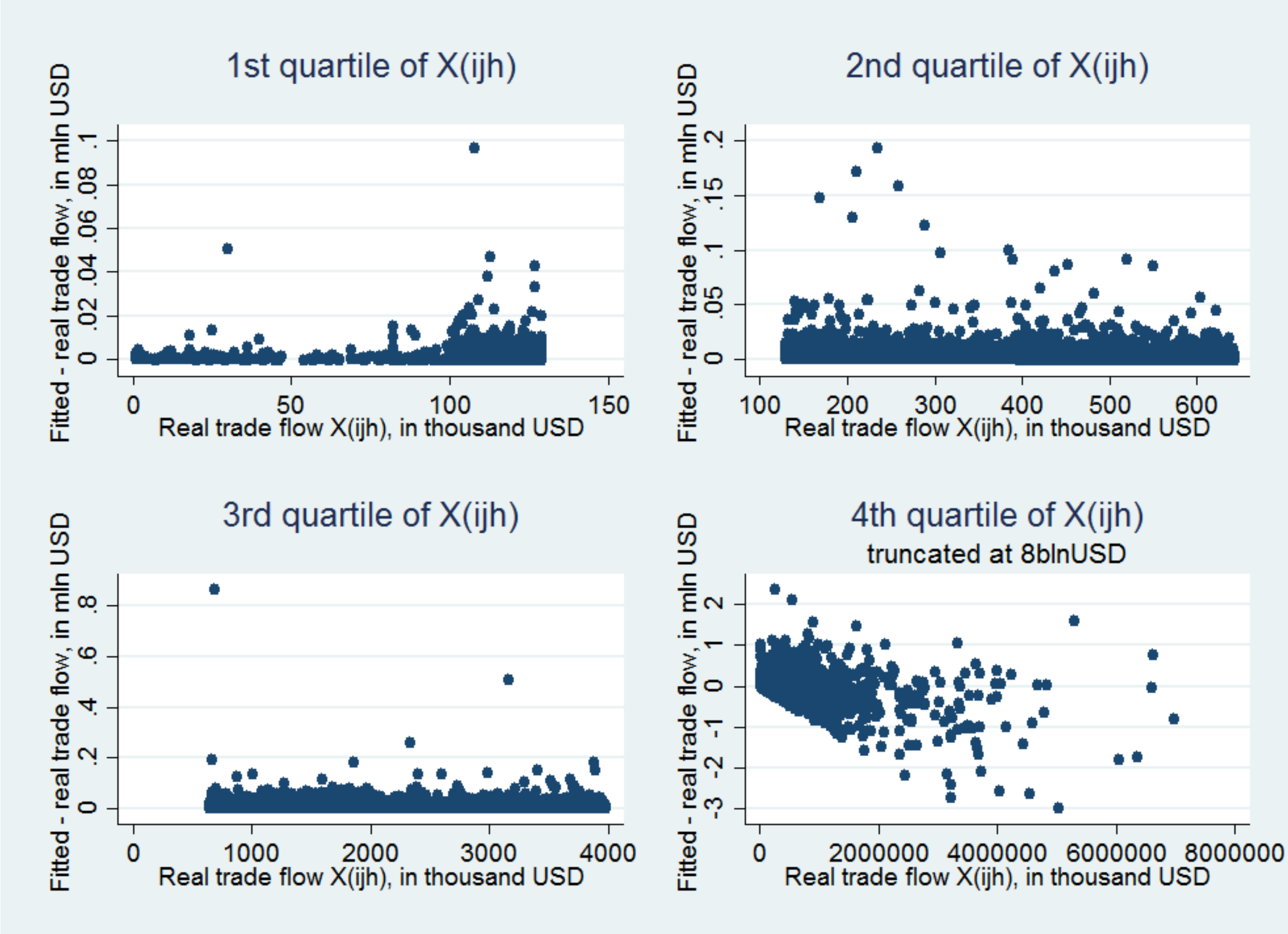
Table A.1 – Distribution of true positive, false positive, true zero and false zero trade flow predicted by the traditional trade gravity equation

		$True\ Value\ (X_{ij}^h > 0)$	
		0	1
$Predicted\ value$ $(\hat{X}_{ij}^h > 0)$	0	48% (2,892,325)	1% (8,139)
	1	52% (3,083,279)	99% (635,713)
	Total	100% (5,975,604)	100% (643,852)

Note: $\hat{X}_{ij}^h$ are the predicted value of $X_{ij} = cost + ln(dist_{ij}) + \theta_i + \theta_j + rta_{ij} + contig_{ij} + comlang_{ij} + col_{ij} + com\_rel_{ij} + com\_cur_{ij} + gat\_memb_i + gat\_memb_j + \varepsilon_{ij}$. Subfix for sector h and time t and parameters notations are omitted for simplification. Parameters of equation are estimated using a Pseudo Poisson Maximum Likelihood (PPML) model, as in Santos Silva and Tenreyro (2007), running separate regressions for each year and sector, on a subsample of the data (years 1981, 1984, 1991, 1992, 1996, 1997) used in Chapter 2. The distribution of the results is extremely stable and similar among years and sectors.

Moreover, the distribution of the fitted values exhibits a particular pattern: the value of bilateral trade is overestimated for low values of real trade flows while is underestimated for high values. In Figure A.1, we plot the difference between the fitted and observed values, conditioned to the value of the observed trade flows. It is straightforward to note that estimates computed with the traditional gravity model tend to underestimate larger observed values while overestimate the smallest. Our hypothesis is that the unobservable component of fixed cost and consumers' preferences lie at the root of the error.

Figure A.1 – Negative error distribution (fitted value minus real value) of the traditional gravity model of trade conditioned to the true trading value (X_{ij}^{h})



Note: This figure shows the distribution of the difference between the fitted values computed from the traditional gravity model of trade and the observed trade flow values. Predicted values are too large for low observed value and small for large observed values. The ratio of bilateral fixed costs to export and consumers' preferences seems to have some intuitive influence on the dynamic of trade.

B Error distribution and potential endogeneity or measurement problem

Considering X_{ij}^h and X_{kj}^h in Equation 24 as stochastic processes with, respectively, multiplicative errors ε_{ijt}^h and ε_{kjt}^h distributed as log-normal variables with mean and standard deviation of the natural logarithm equal to $\mu_{i,k}^h$ and $\sigma_{i,k}^h$, we can write Equation 29 as

$$\frac{X_{ijt}^h}{X_{kjt}^h} = \frac{Y_{it}}{Y_{kt}}\left(\frac{w_{it}\tau_{ijt}^h}{w_{kt}\tau_{kjt}^h}\right)^{-\gamma_h}\left(\frac{\psi_{ij}^h + \psi_{jt}^h + \left(\frac{Y_{st}}{Y_{it}}\left(\frac{w_{st}}{w_{it}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}}\beta^h\sum_{l\neq j}^{L}\frac{\alpha_{slt}^h}{d_{jl}^h}\left[\frac{X_{ilt}^h}{X_{slt}^h}\left(\frac{\tau_{slt}^h}{\tau_{ilt}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h-1}{\gamma_h}}}{\psi_{kj}^h + \psi_{jt}^h + \left(\frac{Y_{st}}{Y_{kt}}\left(\frac{w_{st}}{w_{kt}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}}\beta^h\sum_{l\neq j}^{L}\frac{\alpha_{slt}^h}{d_{jl}^h}\left[\frac{X_{klt}^h}{X_{slt}^h}\left(\frac{\tau_{slt}^h}{\tau_{klt}^h}\right)^{-\gamma_h}\right]^{\frac{\sigma_h-1}{\gamma_h}}}\right)^{\frac{\gamma_h}{\sigma_h-1}}\left(\frac{\varepsilon_{ijt}^h}{\varepsilon_{kjt}^h}\right)$$

(Eq. B.1)

If $ln(\varepsilon_{ijt}^h)\sim N(\mu_i^h,\sigma_{hi}^2)$ and $ln(\varepsilon_{kjt}^h)\sim N(\mu_k^h,\sigma_{hk}^2)$, the random variables ε_{ijt}^h and ε_{kjt}^h are said to have a log-normal distribution with means $M_i^h = exp\left(\mu_i^h + \frac{1}{2}\sigma_{hi}^2\right)$ and $M_k^h = exp\left(\mu_k^h + \frac{1}{2}\sigma_{hk}^2\right)$, respectively. Assuming $ln(\varepsilon_{ijt}^h)$ and $ln(\varepsilon_{kjt}^h)$ are independently distributed $\ln\left(\frac{\varepsilon_{ijt}^h}{\varepsilon_{kjt}^h}\right)$ has mean equal to $\mu_i^h - \mu_k^h + \frac{1}{2}(\sigma_{hi}^2 - \sigma_{hk}^2)$. If $ln(\varepsilon_{ijt}^h)$ and $ln(\varepsilon_{kjt}^h)$ have mean $\mu_i^h$and μ_k^h equal to 0 (are random errors) and $\sigma_{hi}^2 = \sigma_{hk}^2$, $ln\left(\frac{\varepsilon_{ijt}^h}{\varepsilon_{kjt}^h}\right)$ has mean equal to 0 and variance[40] δ_h^2, $ln\left(\frac{\varepsilon_{ijt}^h}{\varepsilon_{kjt}^h}\right)\sim N(0,\delta_h^2)$.

Possible endogeneity measurement error on Equation B.1 can derive from omitted variables captured by ε_{ijt}^h or ε_{ikt}^h that are correlated with X_{ilt}^hor X_{klt}^h respectively and are not captured by the spatially correlated structure of preference, α_{ijt}^h and α_{kjt}^h, neither by any structural parameters of Equation B.1 and are not cancelled out by the ratio of fixed cost $\frac{f_{ijt}^h}{f_{kjt}^h}$. In our opinion, such variables should be variable cost not correctly accounted by $\tau_{ilt,klt}^h$ or productivity shocks (at national or firm level) that are not assumed in the Chaney (2008) model of trade that do not affect the wage level ratio $\frac{w_{it}}{w_{kt}}$. In such cases, $\varepsilon_{ijt,kjt}^h$ should be correlated to $X_{ilt,klt}^h$ through $\varepsilon_{ilt,klt}^h$.

[40] As shown by Zhou (1997), when $n_{i,k}$ are both large, the distribution approximate a standard normal.

$$\frac{X_{ijt}^h}{X_{kjt}^h} = \frac{Y_{it}}{Y_{kt}}\left(\frac{w_{it}\tau_{ijt}^h}{w_{kt}\tau_{kjt}^h}\right)^{-\gamma_h}\left(\frac{\psi_{ij}^h + \psi_{jt}^h + \ \beta^h\left(\frac{Y_{st}}{Y_{it}}\left(\frac{w_{st}}{w_{it}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}}\sum_{l\neq j}^{L}\frac{\alpha_{slt}^h}{d_{jl}^h}\left(\frac{\left(\widehat{X_{ilt}^h}\varepsilon_{ilt}^h\right)\left(\tau_{ilt}^h\right)^{\gamma_h}}{\left(\widehat{X_{slt}^h}\varepsilon_{slt}^h\right)\left(\tau_{slt}^h\right)^{\gamma_h}}\right)^{\frac{\sigma_h-1}{\gamma_h}}}{\psi_{kj}^h + \psi_{jt}^h + \beta^h\left(\frac{Y_{st}}{Y_{kt}}\left(\frac{w_{st}}{w_{kt}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}}\sum_{l\neq j}^{L}\frac{\alpha_{slt}^h}{d_{jl}^h}\left(\frac{\left(\widehat{X_{klt}^h}\varepsilon_{klt}^h\right)\left(\tau_{klt}^h\right)^{\gamma_h}}{\left(\widehat{X_{slt}^h}\varepsilon_{slt}^h\right)\left(\tau_{slt}^h\right)^{\gamma_h}}\right)^{\frac{\sigma_h-1}{\gamma_h}}}\right)^{\frac{\gamma_h}{\sigma_h-1}-1}\left(\frac{\varepsilon_{ijt}^h}{\varepsilon_{kjt}^h}\right)$$

(Eq. B.2)

A solution to this possible measurement error is to compute the fitted value $\hat{X}_{ilt}^h$ and $\hat{X}_{klt}^h$ from the theoretical equation

$$X_{ijt}^h = \ \mu_h \frac{Y_{it}\,Y_{jt}}{Y_t}\left(\frac{w_{it}\tau_{ijt}^h}{\theta_{jt}^h}\right)^{-\gamma_h}\left(\frac{\alpha_{ijt}^h}{f_{ijt}^{h\frac{\gamma_h}{(\sigma_h-1-\gamma_h)}}}\right)^{\frac{\gamma_h}{(\sigma_h-1)}}\varepsilon_{ijt}^h \qquad \text{(Eq. B.3)}$$

and substitute X_{ilt}^h, X_{klt}^h and X_{slt}^h in Equation B.1 with the predicted value $\hat{X}_{ilt}^h, \hat{X}_{klt}^h$ and $\hat{X}_{slt}^h$ computed from Equation B.3.

Log linearizing Equation B.3, we usually control for θ_{jt}^h and Y_{jt} using time varying importer fixed effects, while Y_{it} and w_{it} are absorbed by time varying exporter fixed effects. However, to provide fitted values of trade flows from country *i* to country *j* for sector *h* that are fully compliant with the assumptions of our model, we need to add time varying sector pair country fixed effects, to control for the ratio $\frac{f_{ijt}^h}{\alpha_{ijt}^h}$. The identification of this kind of fixed effect is possible exploiting the product structure of trading data. Assuming $\frac{\alpha_{ijt}^h}{f_{ijt}^{h\frac{\gamma_h}{(\sigma_h-1-\gamma_h)}}}$ and γ_h, σ_h are the same for product *h* belonging to the same upper class *H*, we can add to our sample goods *h+1* that are similar to our target good h belonging to the *H* upper class. It follows that $\left(\frac{\alpha_{ijt}^h}{f_{ijt}^{h\frac{\gamma_h}{(\sigma_h-1-\gamma_h)}}}\right)^{\frac{\gamma_h}{(\sigma_h-1)}} = \left(\frac{\alpha_{ijt}^{h+1}}{f_{ijt}^{h+1\frac{\gamma_{h+1}}{(\sigma_{h+1}-1-\gamma_{h+1})}}}\right)^{\frac{\gamma_{h+1}}{(\sigma_{h+1}-1)}} = \eta_{ijt}^H$. Therefore, Equation B.3 can be estimated using the following Equation B.4.

$$X_{ijt}^{h} = \mu_h \frac{Y_{it}\,Y_{jt}}{Y_t}\left(\frac{w_{it}\tau_{ijt}^{h}}{\theta_{jt}^{h}}\right)^{-\gamma_h}\left(\eta_{ijt}^{H}\right)\varepsilon_{ijt}^{h} \qquad \text{(Eq. B.4)}$$

Exploiting this new sample[41] we are able to identify η_{ijt}^{H} as a time varying fixed effect and compute fitted value of $\hat{X}_{ilt}^{h}, \hat{X}_{klt}^{h}$ and $\hat{X}_{slt}^{h}$ derived directly from the theoretical framework.

Because of the error correlation structure, we have moreover to consider a different distribution for the ratio of log-normal error in Equation B.1. To identify the parameter in equation B.1 we have to solve the system of equation that include equation B.4 and B.2 considering ε_{ijt}^{h} distributed as log-normal variables[42].

Including national productivity shifter in the model is simpler. Following Feenstra et al. (2018), we define A_i the average productivity shifter for country *i*. It follows

$$\frac{X_{ijt}^{h}}{X_{kjt}^{h}} = \frac{Y_{it}}{Y_{kt}}\left(\frac{w_{it}\tau_{ijt}^{h}}{w_{kt}\tau_{kjt}^{h}}\frac{A_{kt}}{A_{it}}\right)^{-\gamma_h}\left(\frac{\psi_{ij}^{h}+\psi_{jt}^{h}+\beta^{h}\left(\frac{Y_{st}}{Y_{it}}\left(\frac{w_{st}A_{it}}{w_{it}A_{st}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}}\sum_{l\neq j}^{L}\frac{\alpha_{slt}^{h}}{d_{jl}^{h}}\left(\frac{\left(\widehat{X_{ilt}^{h}}\varepsilon_{ilt}^{h}\right)\left(\tau_{ilt}^{h}\right)^{\gamma_h}}{\left(\widehat{X_{slt}^{h}}\varepsilon_{slt}^{h}\right)\left(\tau_{slt}^{h}\right)^{\gamma_h}}\right)^{\frac{\sigma_h-1}{\gamma_h}}}{\psi_{kj}^{h}+\psi_{jt}^{h}+\beta^{h}\left(\frac{Y_{st}}{Y_{kt}}\left(\frac{w_{st}A_{kt}}{w_{kt}A_{st}}\right)^{-\gamma_h}\right)^{\frac{\sigma_h-1}{\gamma_h}}\sum_{l\neq j}^{L}\frac{\alpha_{slt}^{h}}{d_{jl}^{h}}\left(\frac{\left(\widehat{X_{klt}^{h}}\varepsilon_{klt}^{h}\right)\left(\tau_{klt}^{h}\right)^{\gamma_h}}{\left(\widehat{X_{slt}^{h}}\varepsilon_{slt}^{h}\right)\left(\tau_{slt}^{h}\right)^{\gamma_h}}\right)^{\frac{\sigma_h-1}{\gamma_h}}}\right)^{\frac{\gamma_h}{\sigma_h-1}-1}\left(\frac{\varepsilon_{ijt}^{h}}{\varepsilon_{kjt}^{h}}\right)$$

---

[41] For example, to compute time varying fixed effect for the HS 6101 class product (Men's coats) we can add to our sample products HS 6102 (Girl's coats), HS 6103 (Men's Suits), HS 6104 (Women's suits) and so on. Using a higher detail (HS at 6 digit) permits to consider a higher number of similar products.

[42] Value of ε_{ijt}^{h} and $\widehat{X_{ilt}^{h}}$ can be computed or sampled using Equation B.4, specifying a particular spatial correlation structure of ε_{ijt}^{h}.

C Estimation of Equation 30 on simulated data

To assess the performance of our identification and estimation strategy, we compute the parameters of Equation 30 on some simulated datasets. Number of countries (*i*,*j*,*k*,*l*) and time periods (*t*) are equals to 10. ψ_{ij}^h is distributed in the domain [0,1], ψ_{jt}^h is equal to 0.1, 0.15, 0.05, 0.06, 0.07, 0.09, 0.12, 0.11, 0.13, 0.08 for *j*=1,...,10 and *t* equal to 1 while, for *t* greater than 1, ψ_{jt}^h is equal to $\psi_{jt-1}^h + N(0,1)/100$. α_{slt}^h is equal to 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.25, 0.85, 0.95 for *t* = 1 and *l*=1,…,10, while, for *t* greater than 2, is equal to $\alpha_{slt-1}^h + N(0,1)/10$. τ_{ijt}^h, Y_{it} and w_{it} are uniformly distributed in the domain [1.1, 1.15], [150000, 3580000] and [358, 1500] respectively. Finally, δ and d_{jl}^h are set equal to one for all *j*,*l*.

To estimate model's parameters according to the theoretical framework, we use the following objective priors:

$\sigma = 1 + \sigma_1$

$\gamma = 1 + \sigma_1 + \gamma_1$

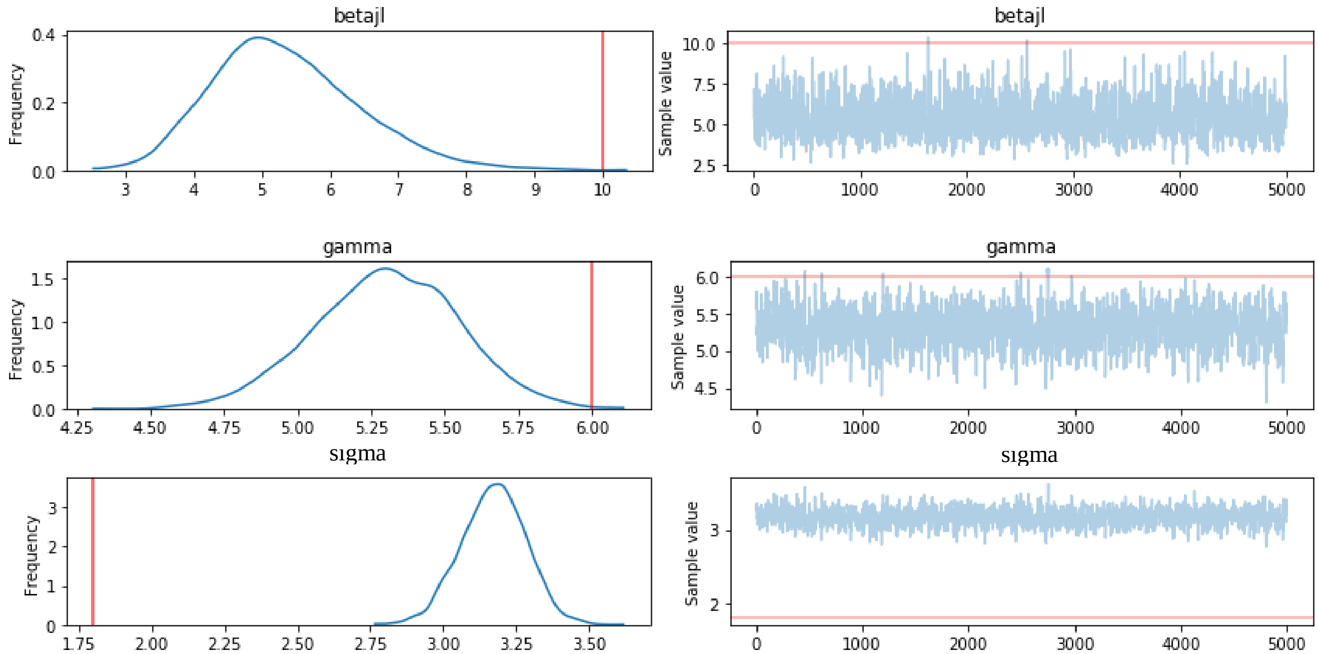
$\psi_{ij}^h, \psi_{jt}^h, \alpha_{sl}^h, \sigma_1, \gamma_1 \sim \exp(1.5)$

$\delta \sim HalfCauchy(5)$

All parameters are sampled using a No-U-Turn sampler (Hoffman and Gelman, 2014) and estimates are computed using 10000 draws, discarding the first 5000.

For the first simulation, γ_h and σ_h are set equal to 6 and 1.8 respectively while β^h is equal to 10. Results are in Figure B.1.

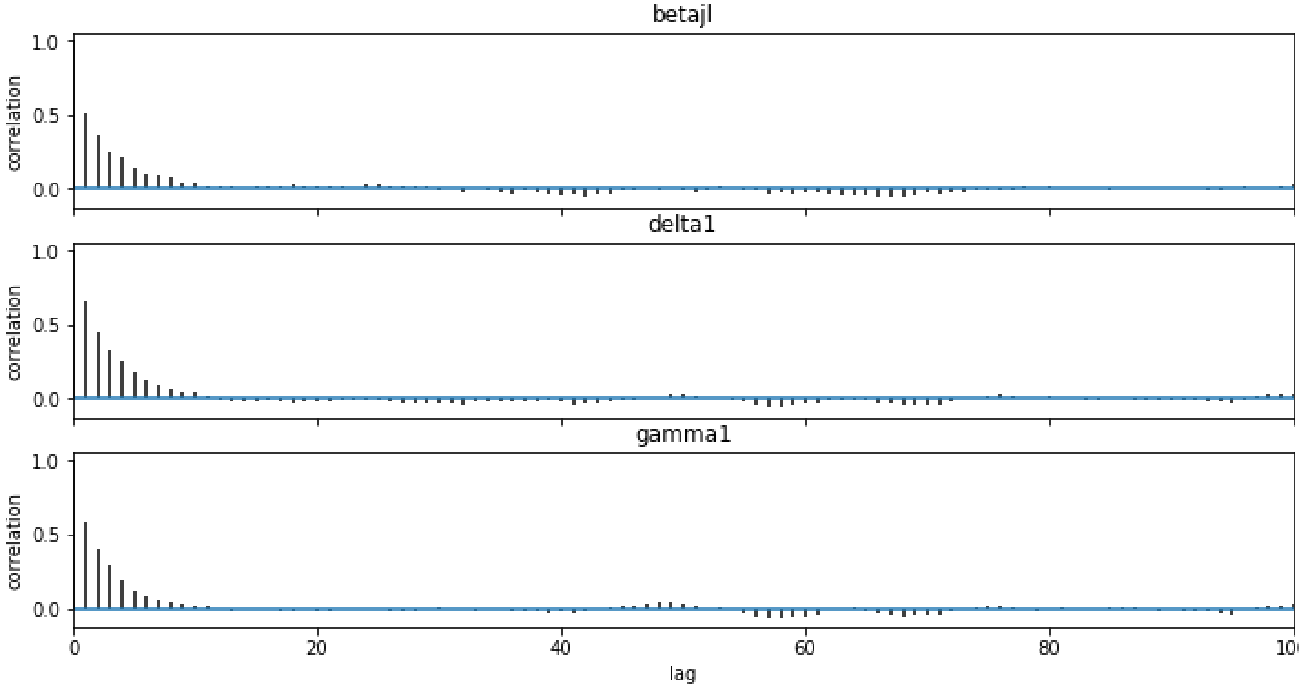
Figure C.1 – Posterior means and traces of parameters β^h, γ_h and σ_h for the simulated dataset, with true values 10, 6 and 1.8.



Note: In Figure b.1 we observe the distribution of the estimated spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 30 for a simulated dataset. Values are computed using Hamiltonian Monte Carlo method, with a target acceptance rate of 0.95 for the No-U-Turn sampler and 10.000 draws (with 5.000 burned draws). Red lines are true values of the parameters in the simulated dataset.

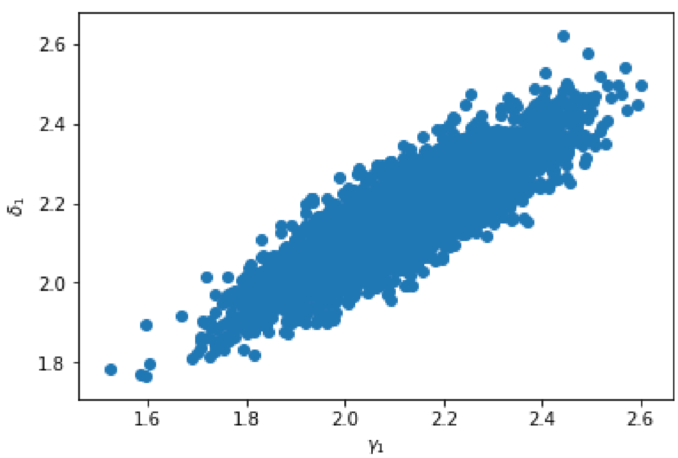
It is easy to show that estimates for β^h and γ_h are downward biased while σ_h is overestimated. A visual inspection of autocorrelation in the trace of parameter (Figure B.2) does not exhibit a high level of correlation, after discarding 5.000 draws.

Figure C.2 – Autocorrelation plot of the posterior means for parameters β^h, σ_1 and γ_1 for the simulated dataset with true values 10, 1.8 and 6.



Note: In Figure C.2 we observe the autocorrelation plot of a subset of parameters of Equation 30. Values are computed using Hamiltonian Monte Carlo computation, with a target acceptance rate of 0.95 for the No-U-Turn sampler and 10.000 draws (with 5.000 burned draws).

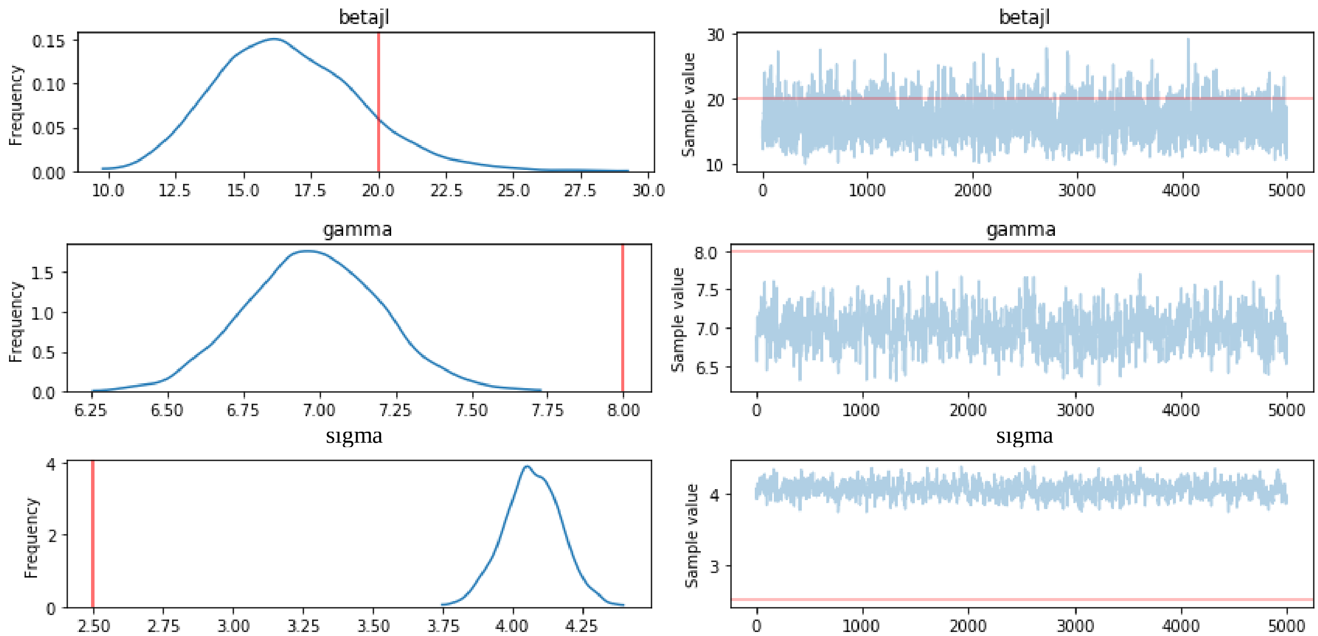
Figure C.3 – Distribution of the posterior means of σ_1 and γ_1 for the simulated dataset



Note: In figure C.3 we observe the jointed distribution of posterior means for σ_1 and γ_1 for the simulated dataset.

Simulating a new dataset with higher values of γ_h, σ_h and β^h, equal respectively to 8, 2.5 and 20, the bias in absolute value for estimates of γ_h and σ_h increases while β^h accuracy improves. The spatial correlation parameter β^h is now included in the 95% confidence interval. With this new simulated dataset, the true posterior mean is about 14% higher for γ_h, 38% lower for σ_h and 20% higher for β^h. From a relative point of view, higher true parameters' values gives best posterior estimates.

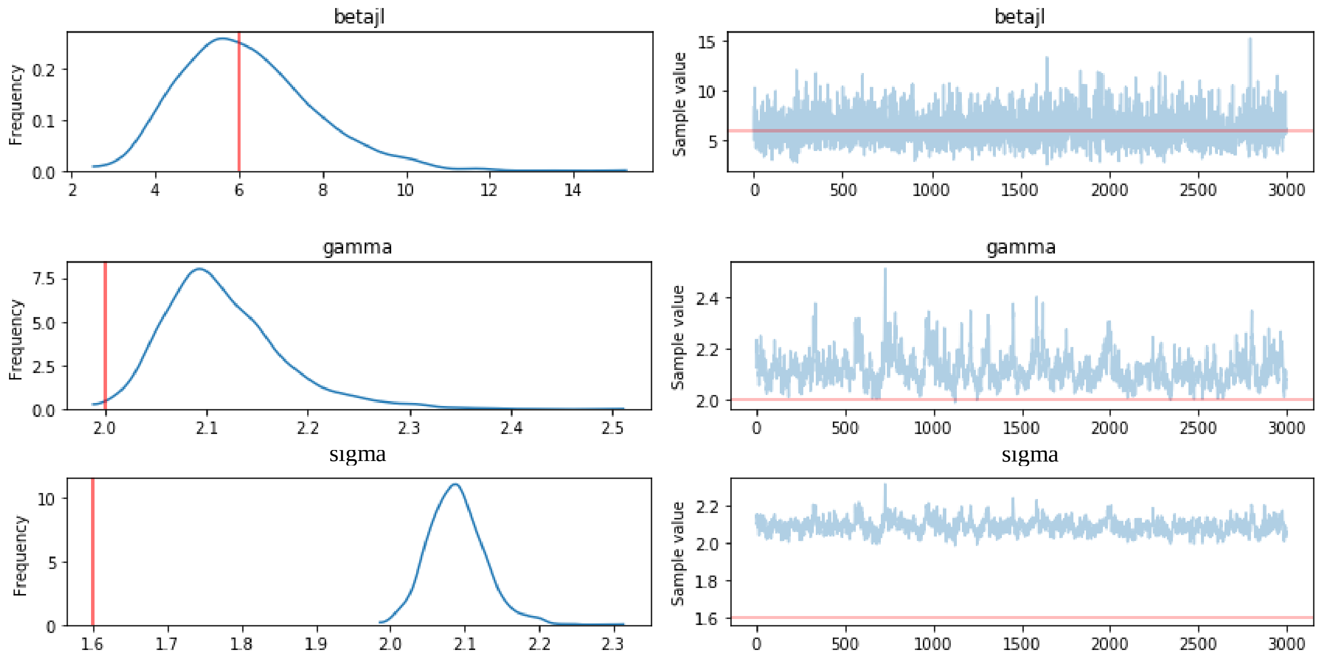
Figure C.4 – Posterior means and traces of parameters β^h, γ_h and σ_h for the simulated dataset, with true values 20, 8 and 2.5.



Note: In this figure we observe the distribution of the estimated spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 30. Values are computed using Hamiltonian Monte Carlo method, with a target acceptance rate of 0.95 for the No-U-Turn sampler and 10.000 draws (with 5.000 burned draws). Red lines are true values of the parameters in the simulated dataset.

Using a third simulated dataset with values close to our estimates (β^h, γ_h, and σ_h equal to 6, 2 and 1.6 respectively), the posterior means of the parameters tend to be closer to their true values. The coefficient of the spatial lag, β^h, is close to its exact value and even the other parameters exhibit lower bias (Figure B.5).

Figure C.5 – Posterior means and traces of parameters β^h, γ_h and σ_h for the simulated dataset, with true values 6, 2 and 1.6.



Note: In this figure we observe the distribution of the estimated spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 30. Values are computed using Hamiltonian Monte Carlo method, with a target acceptance rate of 0.95 for the No-U-Turn sampler and 6.000 draws (with 3.000 burned draws). Red lines are true values of the parameters in the simulated dataset.

To conclude, estimates of β^h tend to be within the confidence interval or, in the worst case, are downward bias. Given that the hypothesis we test in this paper is $\beta^h > 0$, because we suppose preferences are spatially correlated, the downward bias does not affect negatively our conclusion. However, at this moment, we cannot affirm that the posterior means of σ_h and γ_h are the true elasticity of substitution and heterogeneity distribution parameter of good h.

D Estimates of Equation 30 in Chapter 4.2

In order to estimate the parameters of Equation 30, we use the following objective priors:

$\sigma = 1 + \sigma_1$

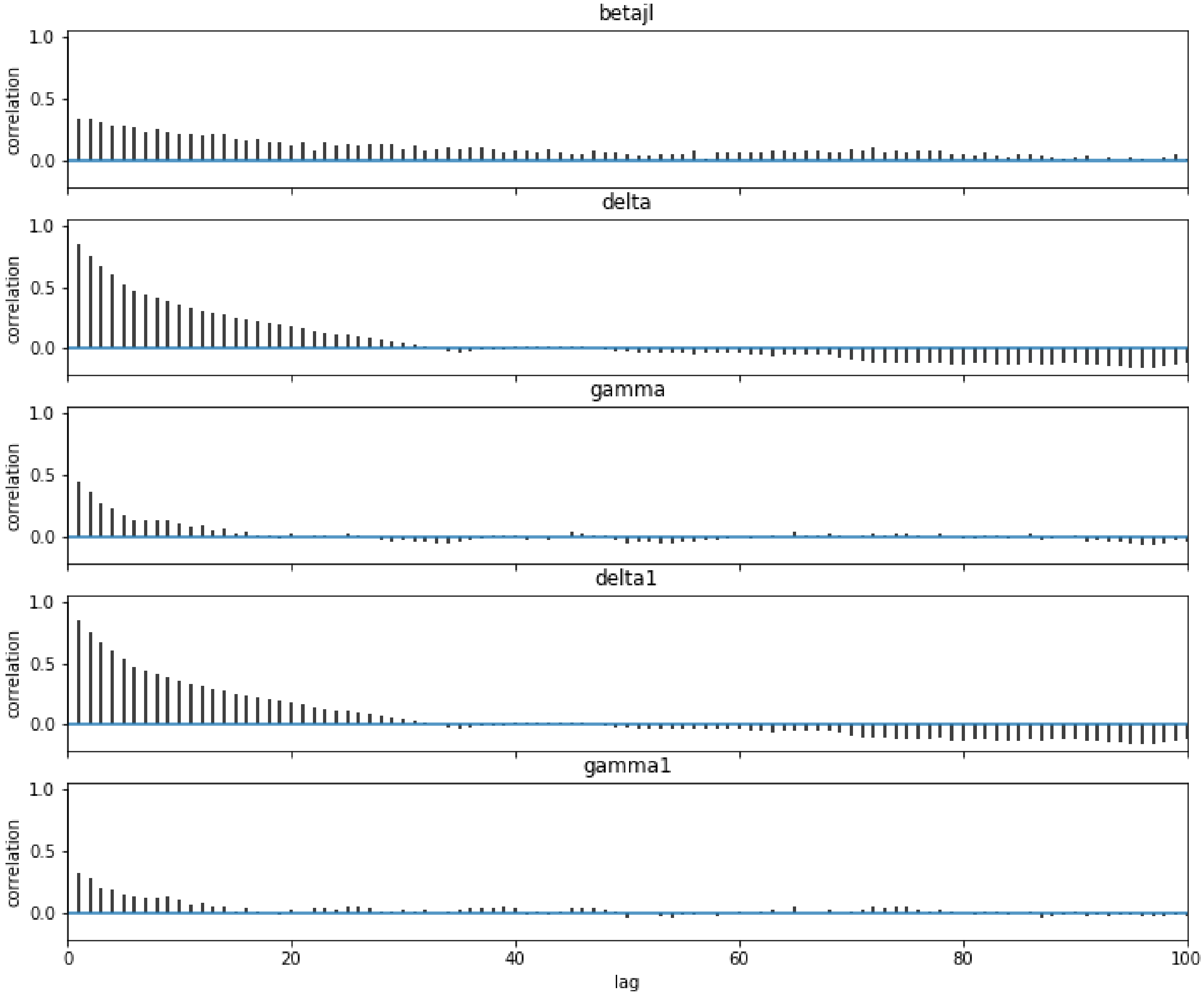
$\gamma = 1 + \sigma_1 + \gamma_1$

$\psi_{ij}^h, \psi_{jt}^h, \alpha_{slt}^h, \sigma_1 \sim \exp(1.5)$

$\gamma_1 \sim \exp(0.8)$

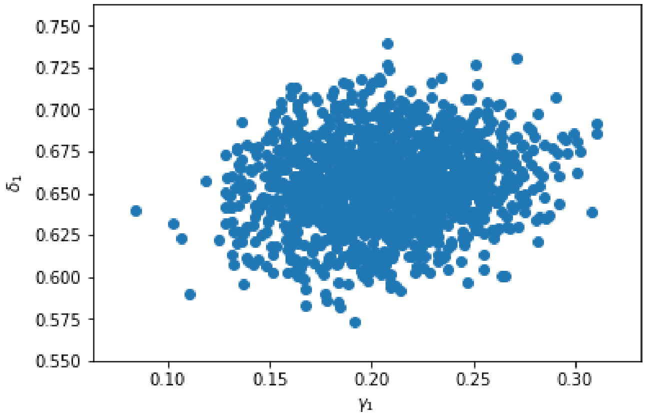
The posterior mean of parameters are computed on the value estimated using the Hoffman and Gelman (2014) No-U-Turn sampler. Results and diagnostic are in the following figures, for the subset of sectors.

Figure D.1 – Autocorrelation plot of the posterior means for parameters β^h, σ_1, γ_1, σ_1, γ_h, σ_h for product 22 HS (beverages, spirits and vinegar)



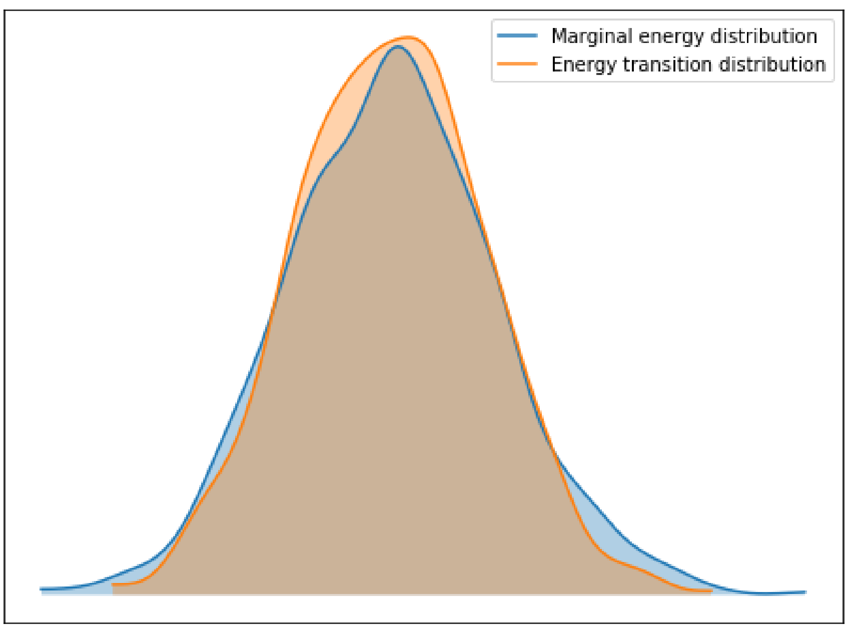
Note: In this figure we observe the autocorrelation plot of a subset of parameters of Equation 30. Values are estimates using Hamiltonian Monte Carlo computation, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.2 – Distribution of the posterior means of σ_1 and γ_1 for product 22 HS (beverages, spirits and vinegar)



Note: This figure shows the joint posterior means distribution of σ_1 and γ_1.

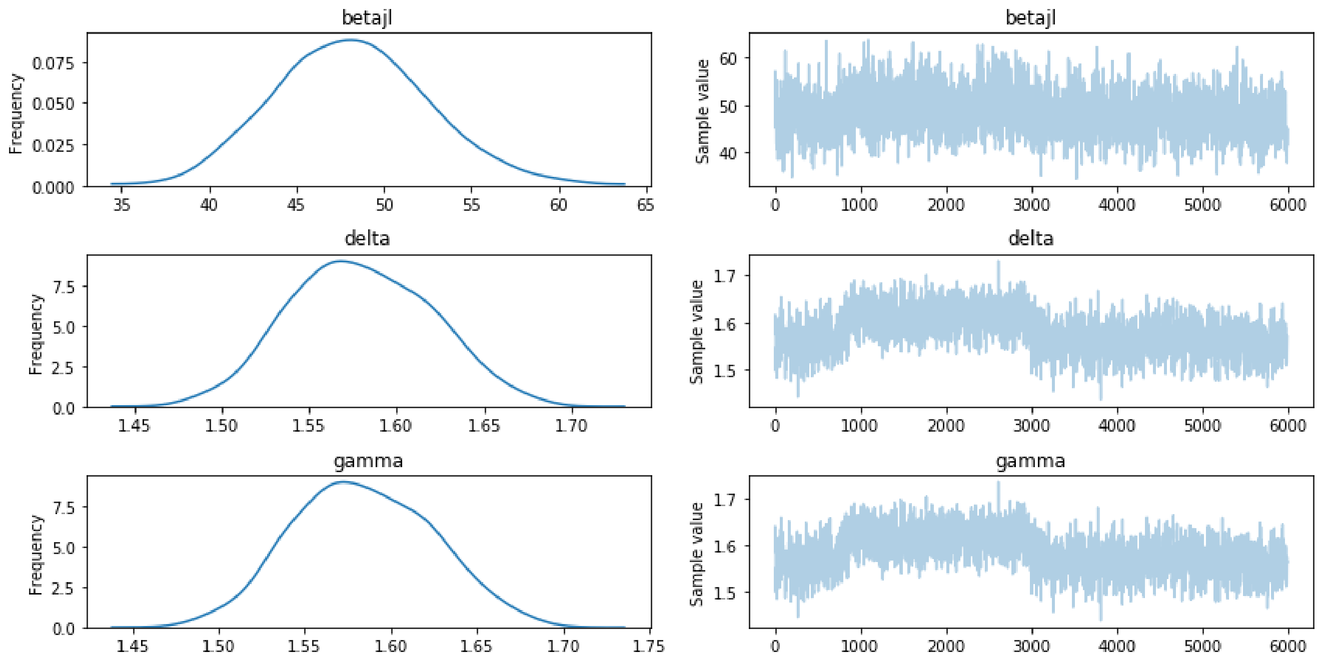
Figure D.3 – Energy plot distribution for product 22 HS (beverages, spirits and vinegar)



Note: Plot energy transition distribution and marginal energy distribution in order to diagnose poor exploration by HMC algorithm.

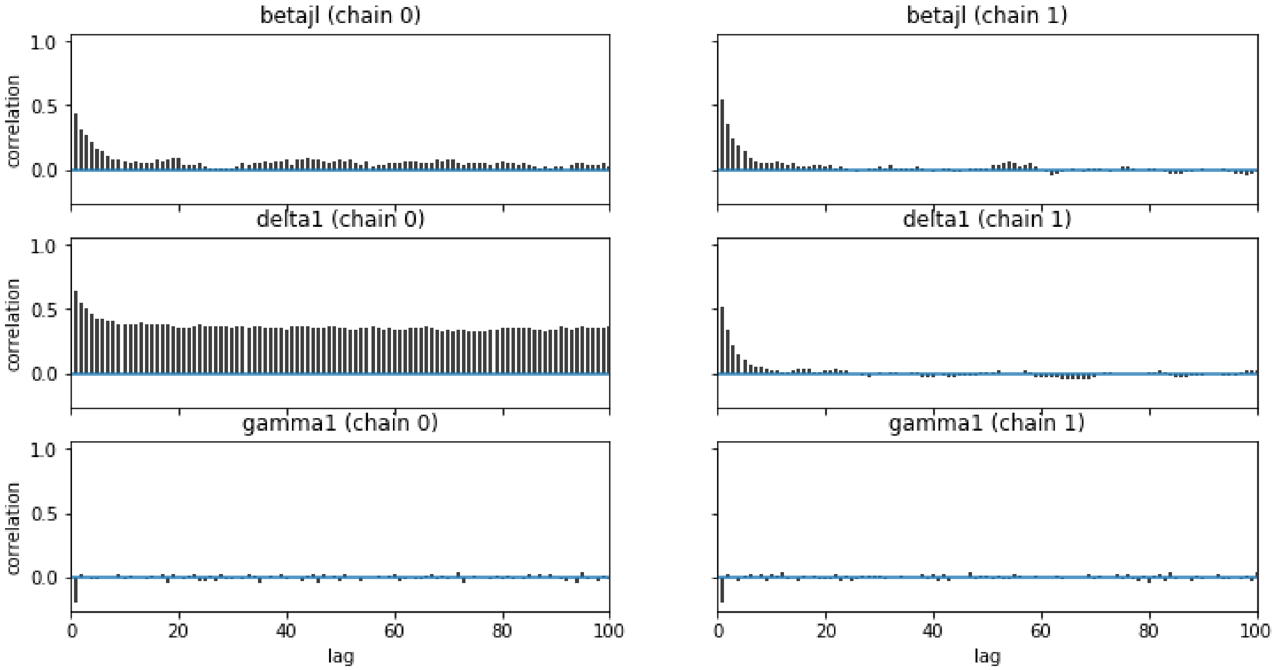
Figure D.4 – Posterior means and traces of parameters β^h, σ_h and γ_h of Table 3, for the HS 61 Sector (Apparel and clothing accessories; knitted or crocheted)

Combined Graphs



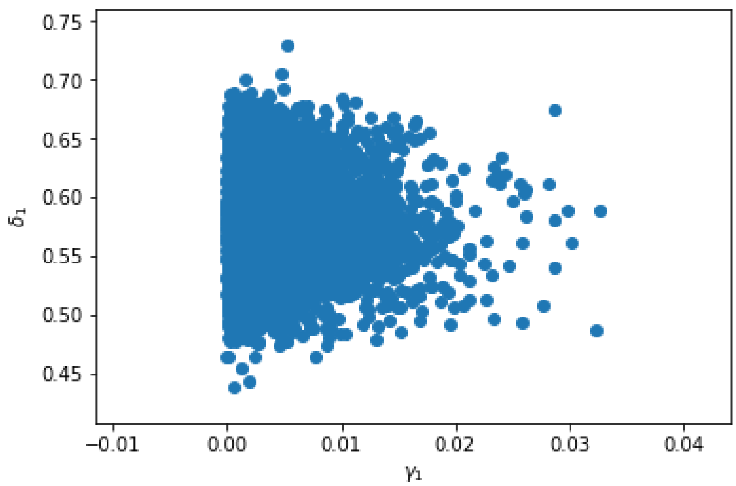
Note: This figure displays the distribution of the estimated spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 30. Values are computed using Hamiltonian Monte Carlo method, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.5 – Autocorrelation plot of the posterior means for parameters $\beta^h, \sigma_1, \gamma_1, \sigma_1, \gamma_h, \sigma_h$, for the HS 61 Sector (Apparel and clothing accessories; knitted or crocheted)



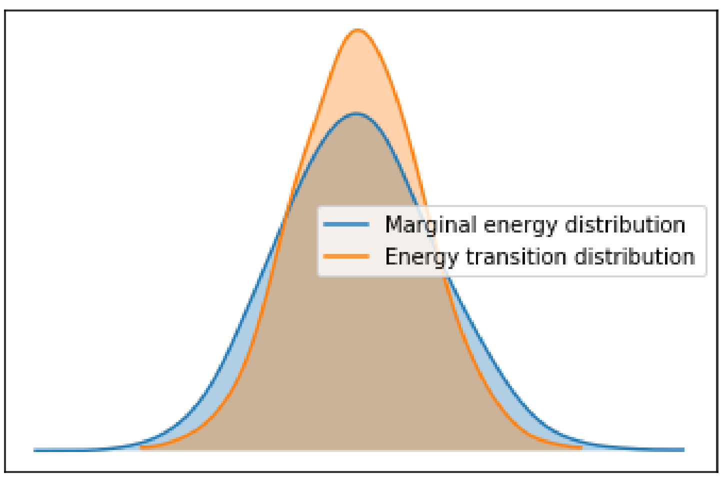
Note: In this figure we observe the autocorrelation plot of a subset of parameters of Equation 30. Values are estimates using Hamiltonian Monte Carlo computation, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.6 – Distribution of the posterior means of σ_1 and γ_1 for the HS 61 Sector (Apparel and clothing accessories; knitted or crocheted)



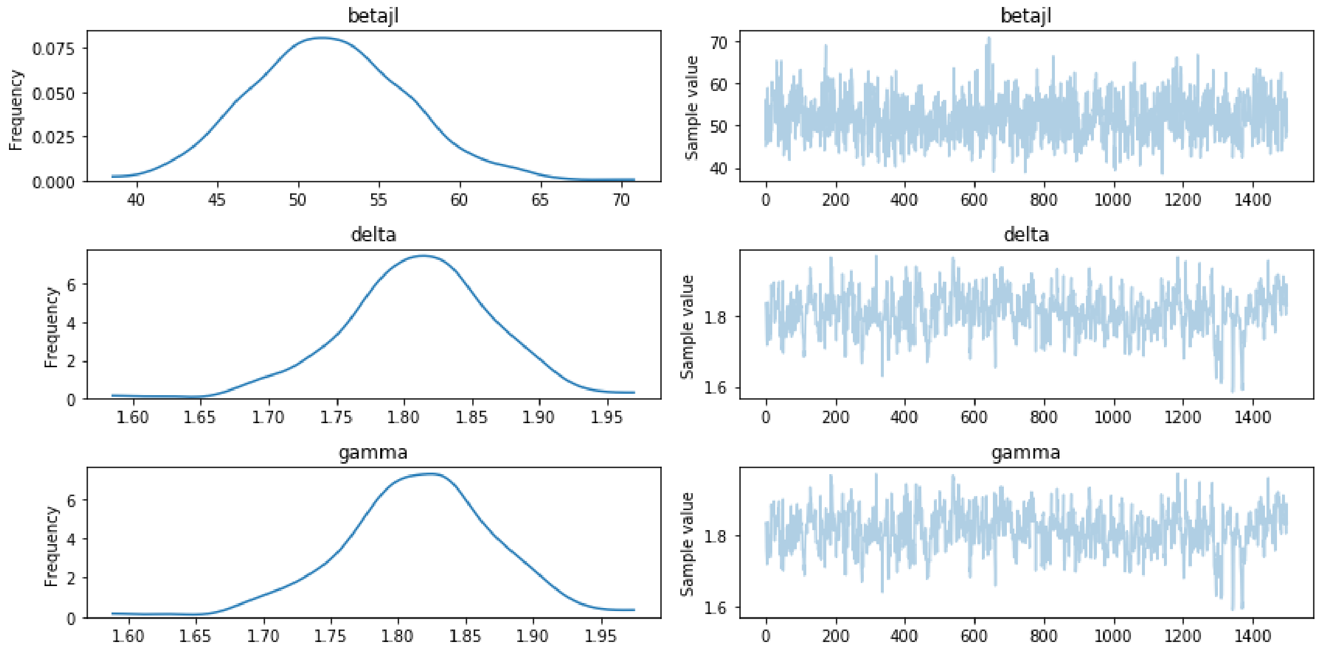
Note: This figure shows the joint posterior mean distribution of σ_1 and γ_1.

Figure D.7 – Energy plot distribution for the HS 61 Sector (Apparel and clothing accessories; knitted or crocheted)



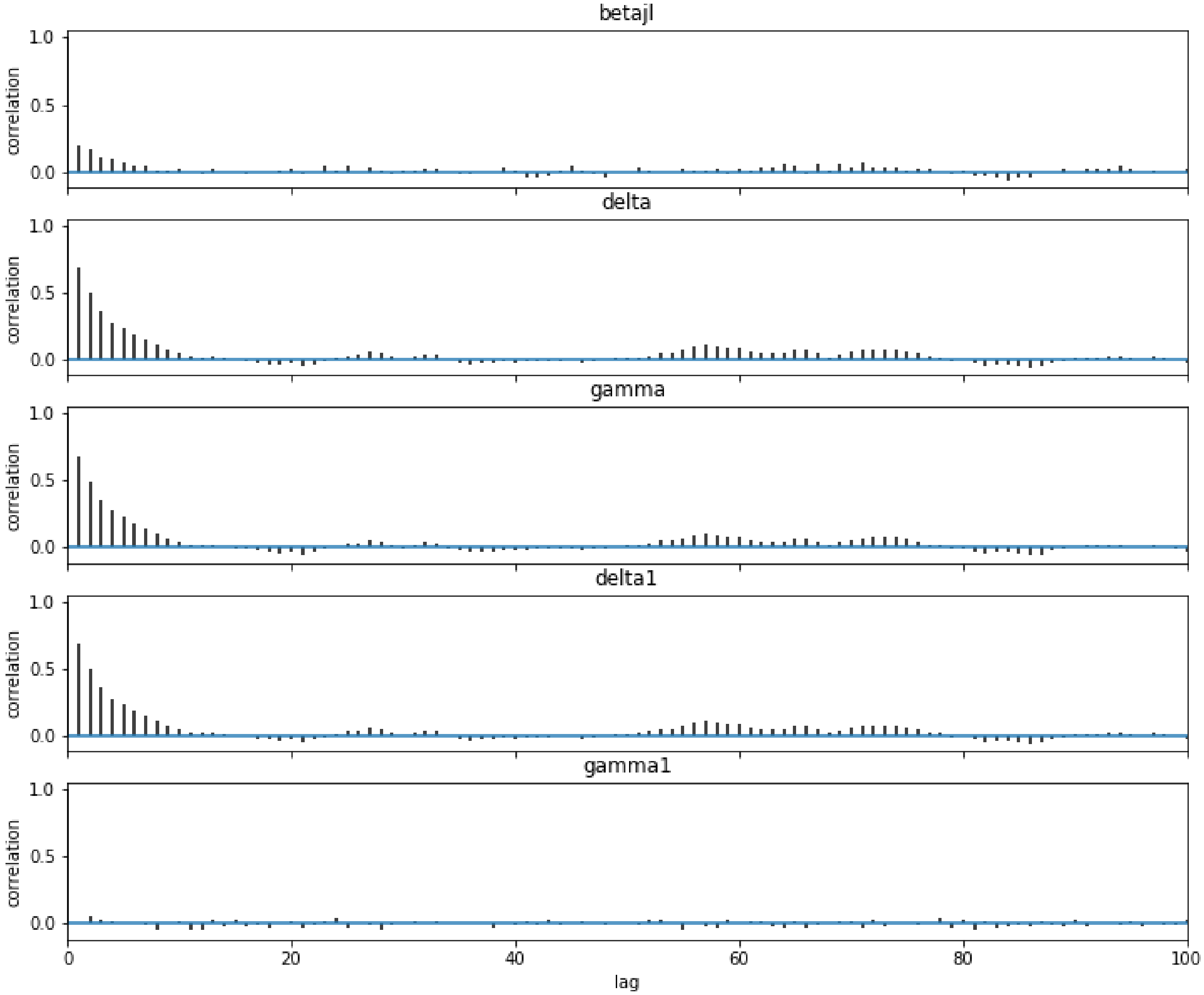
Note: Plot energy transition distribution and marginal energy distribution in order to diagnose poor exploration by HMC algorithm.

Figure D.8 – Posterior means and traces of parameters β^h, σ_h and γ_h of Table 3, for the HS 38 Sector (Miscellaneous chemical products)



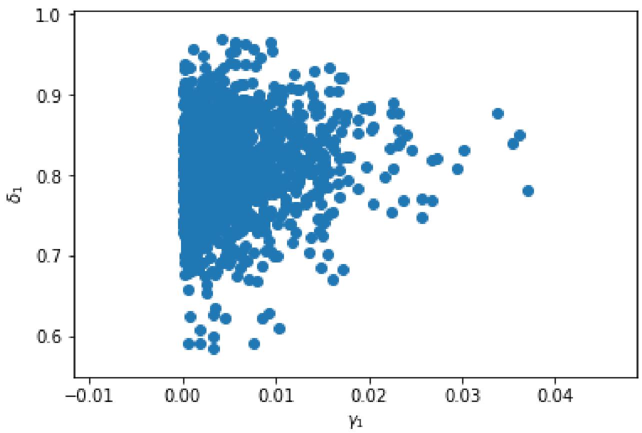
Note: This figure displays the distribution of the estimated spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 30. Values are computed using Hamiltonian Monte Carlo method, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.9 – Autocorrelation plot of the posterior means for parameters $\beta^h, \sigma_1, \gamma_1, \sigma_1, \gamma_h, \sigma_h$, for the HS 38 Sector (Miscellaneous chemical products)



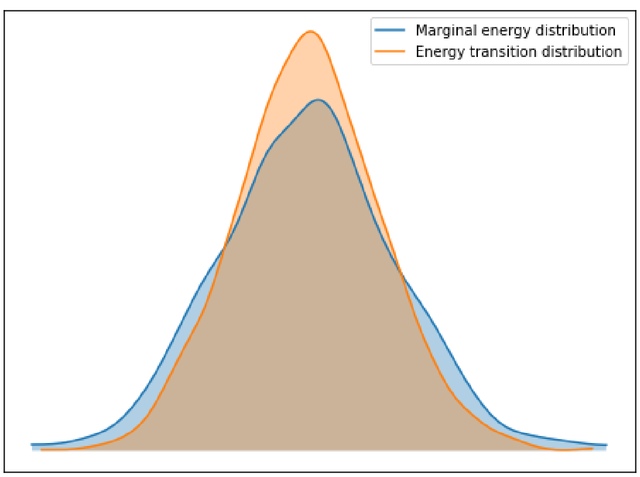
Note: In this figure we observe the autocorrelation plot of a subset of parameters of Equation 30. Values are estimates using Hamiltonian Monte Carlo computation, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.10 – Distribution of the posterior means of σ_1 and γ_1 for the HS 38 Sector (Miscellaneous chemical products)



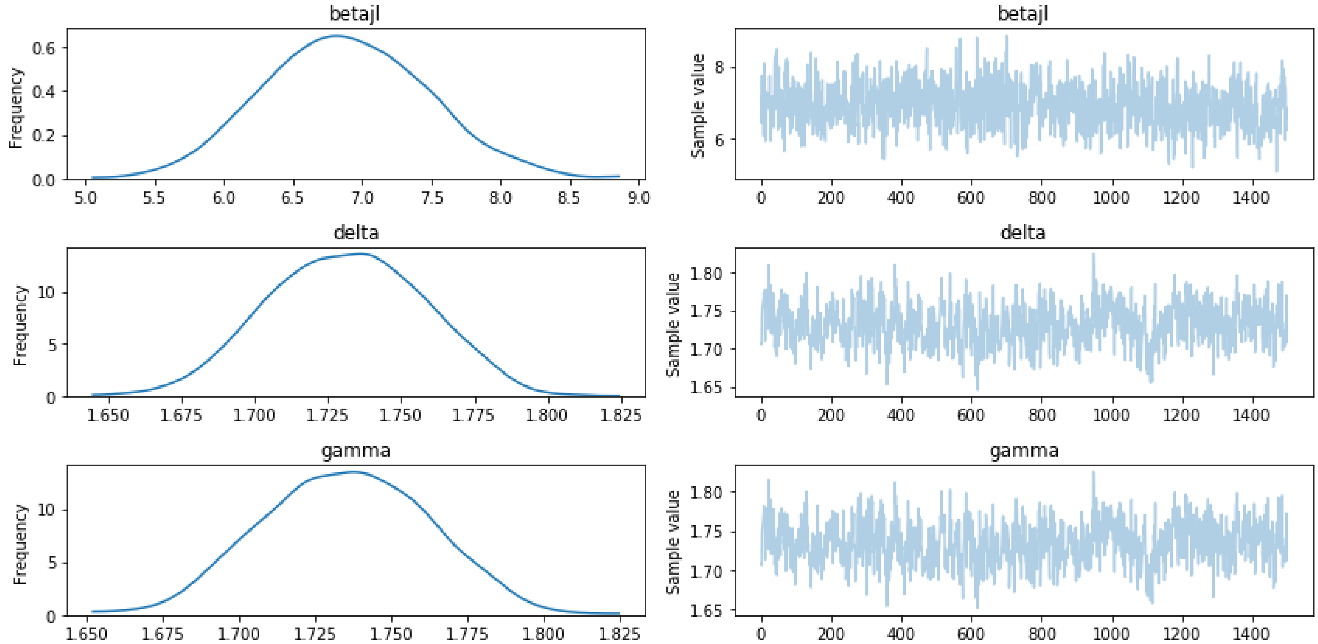
Note: This figure shows the joint posterior mean distribution of σ_1 and γ_1.

Figure D.11 – Energy plot distribution for the HS 38 Sector (Miscellaneous chemical products)



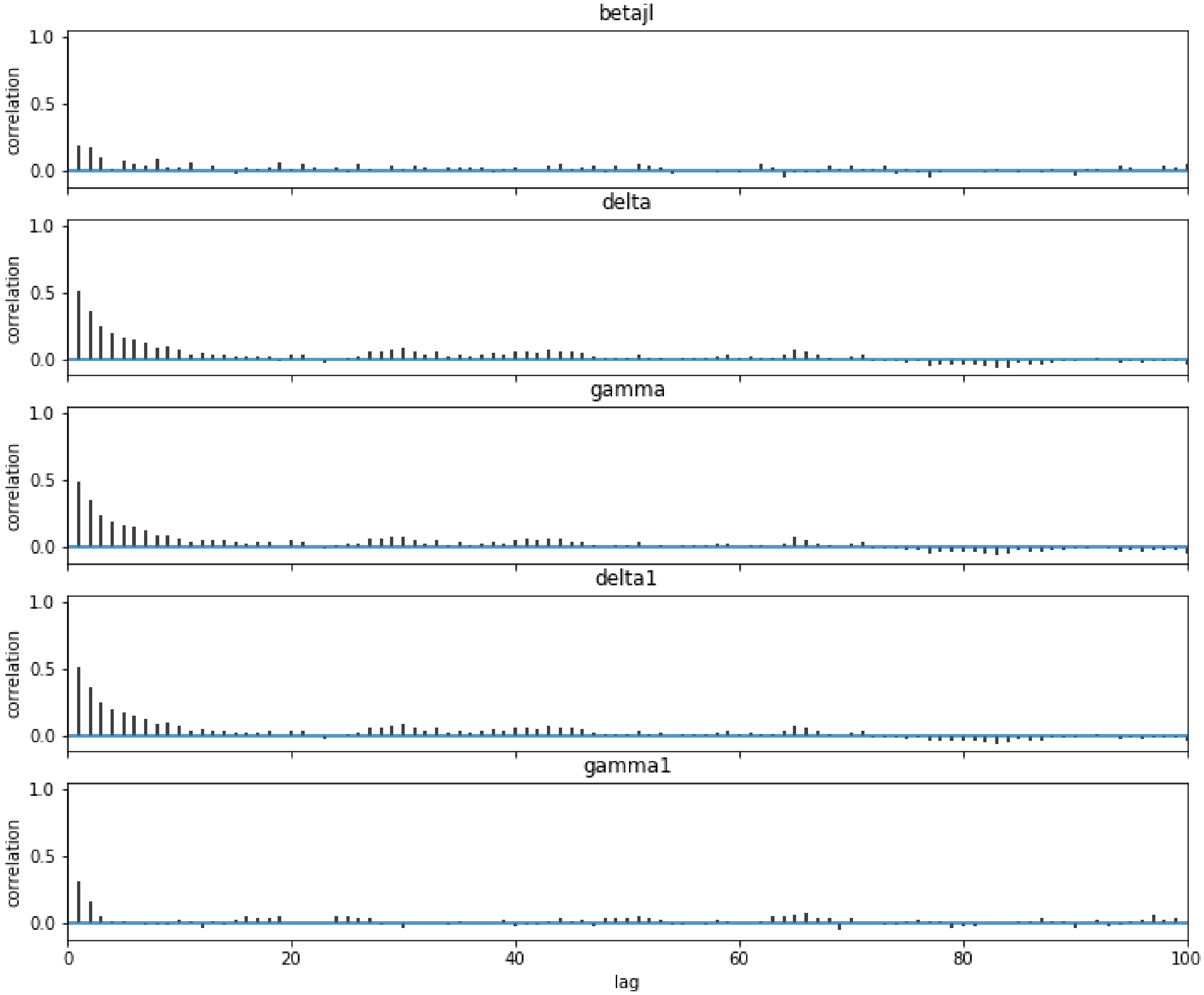
Note: Plot energy transition distribution and marginal energy distribution in order to diagnose poor exploration by HMC algorithms.

Figure D.12 – Posterior means and traces of parameters β^h, σ_h and γ_h of Table 3, for the HS 69 Sector (Ceramic products).



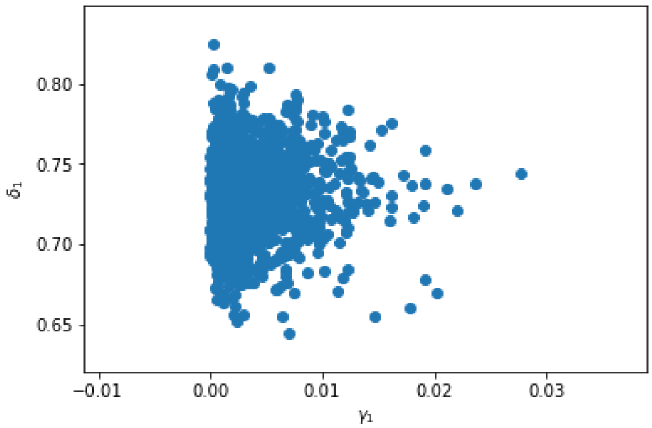
Note: This figure displays the distribution of the estimated spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 30. Values are computed using Hamiltonian Monte Carlo method, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.13 – Autocorrelation plot of the posterior means for parameters β^h, σ_1, γ_1, σ_1, γ_h, σ_h, for the HS 69 Sector (Ceramic products).



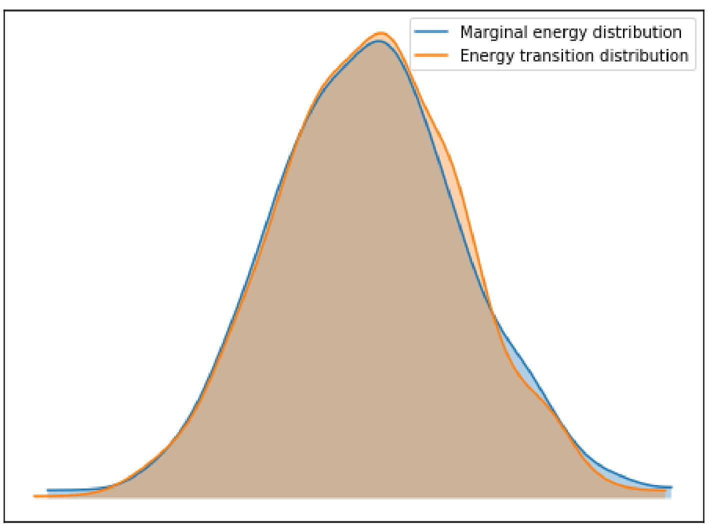
Note: In this figure we observe the autocorrelation plot of a subset of parameters of Equation 30. Values are estimates using Hamiltonian Monte Carlo computation, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.14 – Distribution of the posterior means of σ_1 and γ_1 for the HS 69 Sector (Ceramic products).



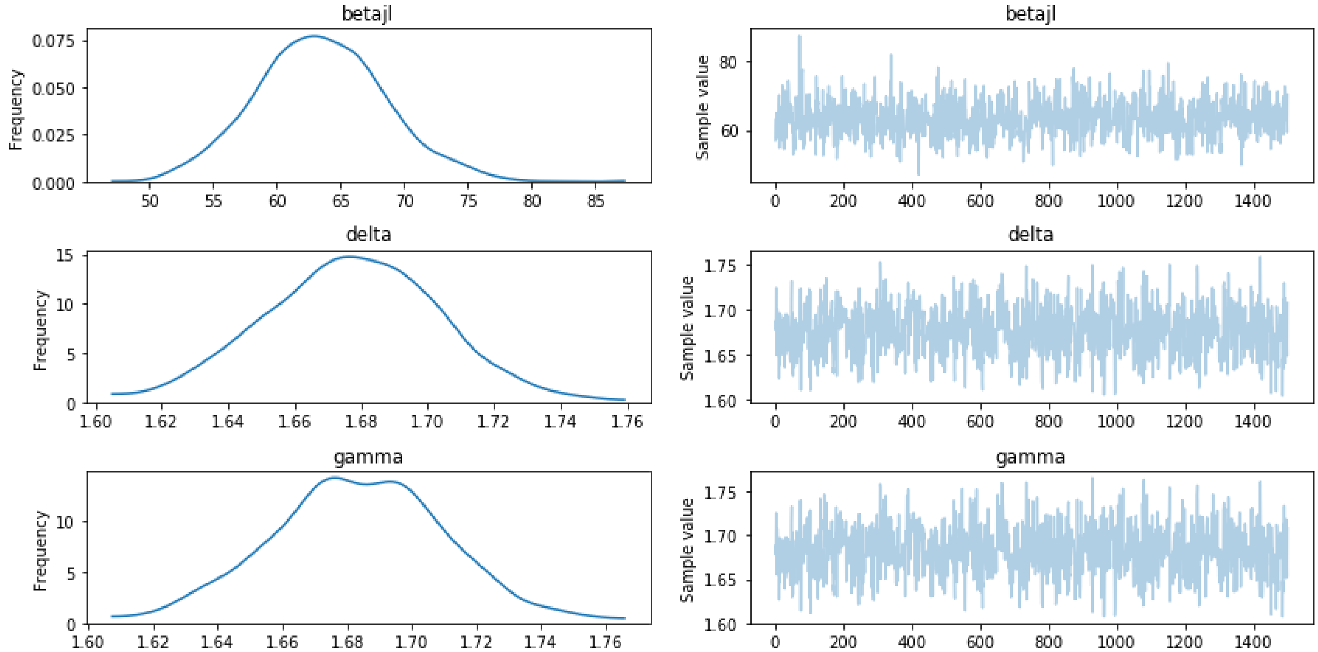
Note: This figure shows the joint posterior mean distribution of σ_1 and γ_1.

Figure D.15 – Energy plot distribution for the HS 69 Sector (Ceramic products).



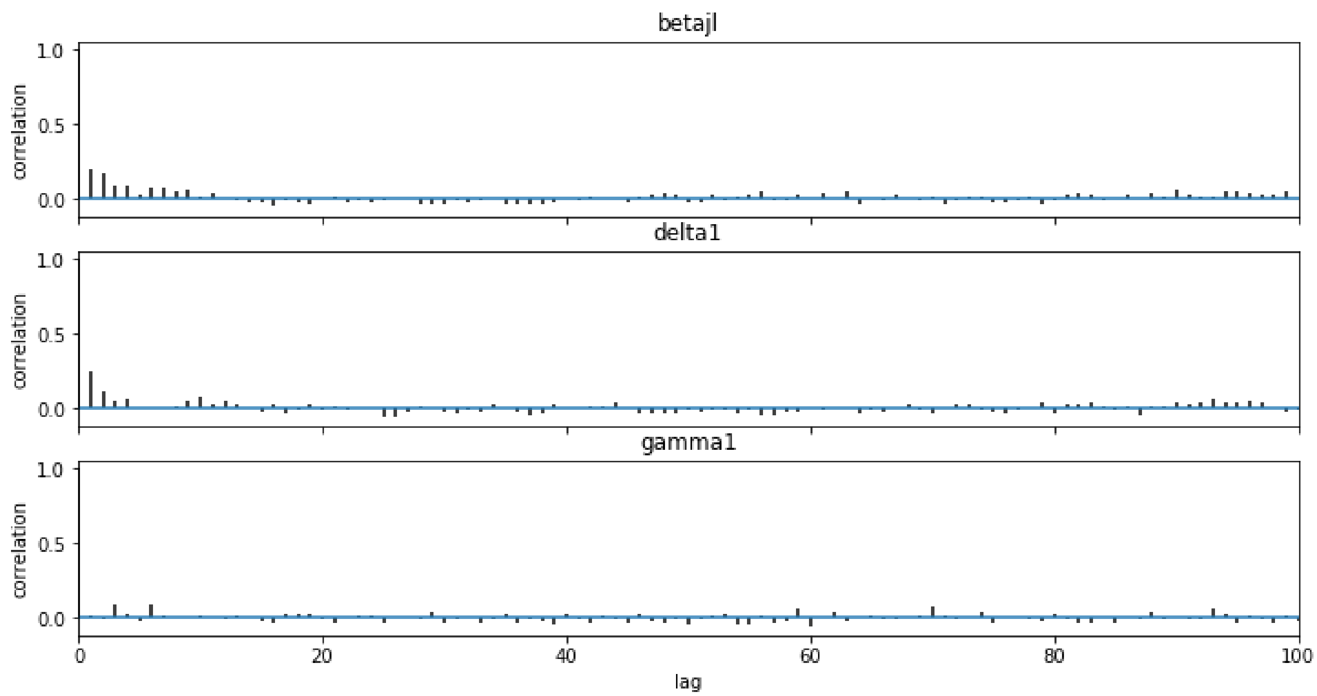
Note: Plot energy transition distribution and marginal energy distribution in order to diagnose poor exploration by HMC algorithms.

Figure D.16 – Posterior means and traces of parameters β^h, σ_h and γ_h of Table 3, for the HS 87 Sector (Vehicles; other than railway or tramway rolling stock).



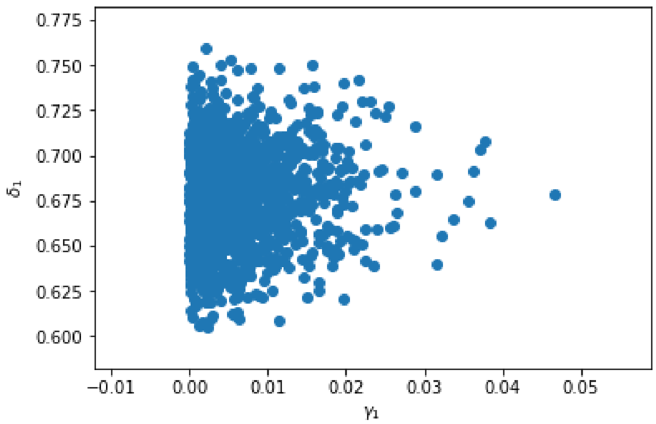
Note: This figure displays the distribution of the estimated spatial correlation parameter (β^h), elasticity of substitution (σ_h) and productivity heterogeneity parameter (γ_h) of Equation 30. Values are computed using Hamiltonian Monte Carlo method, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.17 – Autocorrelation plot of the posterior means for parameters $\beta^h, \sigma_1, \gamma_1, \sigma_1, \gamma_h, \sigma_h$, for the HS 87 Sector (Vehicles; other than railway or tramway rolling stock).



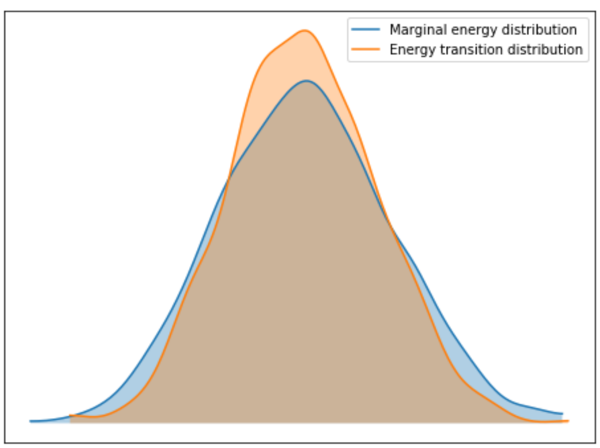
Note: In this figure we observe the autocorrelation plot of a subset of parameters of Equation 30. Values are estimates using Hamiltonian Monte Carlo computation, with a target acceptance rate of 0.95 for the No-U-Turn sampler.

Figure D.18 – Distribution of the posterior means of σ_1 and γ_1 for the HS 87 Sector (Vehicles; other than railway or tramway rolling stock).



Note: this figure show the joint posterior mean distribution of σ_1 and γ_1.

FigureD.19 – Energy plot distribution for the HS 87 Sector (Vehicles; other than railway or tramway rolling stock).



Note: Plot energy transition distribution and marginal energy distribution in order to diagnose poor exploration by HMC algorithms.

E Non normalized and row-normalized distance matrices

In order to give more interpretable results, we try to estimate our spatial correlation parameter in a more formal way, using the standard matrix approach. Adding an error term to Equation 17 and rewriting in a matrix notation we have:

$$A_i^h = (I - \beta^h W)^{-1} Z_i^h + (I - \beta^h W)^{-1} \varepsilon_i^h \qquad \text{(d.1)}$$

where A_i^h is the vector of preferences α_{ij}^h, I is the identity matrix of dimension *n* (where *n* is the number of countries in the custom union), β^h the spatial correlation parameter, W the inverse distance matrix whose elements are $\frac{1}{d_{jk}}$ before row-normalizing, Z_i^h the vector of idiosyncratic preferences $z_{ij}^h \phi$, and ε_i^h is the vector of error terms ε_{ij}^h distributed as a normal with zero mean, variance δ_{ij}^2, and $cov(\varepsilon_{ij}^h, \varepsilon_{ik}^h) = 0$.

The last term of Equation d.1 can therefore be written as a multivariate Normal distributed variable $\xi \sim N_j(0, \Sigma)$ where $\Sigma = (I - \beta^h W)^{-1} \varepsilon (I - \beta^h W')^{-1}$ and ε is a square matrix of dimension *J* with $\varepsilon_{jj} = \delta_{ij}^2$, $\varepsilon_{jk} = 0$. More generally, Equation d.1 can be written as a multivariate Normal distribution $A_i^h \sim N_j\big((I - \beta^h W)^{-1} Z_i^h;\ (I - \beta^h W)^{-1} \varepsilon (I - \beta^h W')^{-1}\ \big)$.

Adding the time dimension *t*, we have *t*i* vectors of A_{it}^h and Z_{it}^h for each product *h*. Our preference parameters α_{ijt}^h are therefore computed from the multivariate Gaussian distributions $A_{it}^h = N_{jt}\big((I - \beta^h W)^{-1} Z_{it}^h;\ (I - \beta^h W)^{-1} \varepsilon (I - \beta^h W')^{-1}\ \big)$ where β^h, ε are unknown parameters and Z_{it}^h are vectors of unobservable idiosyncratic preferences z_{ijt}^h, equal to:

$$z_{ijt}^h = \psi_{ij}^h + \psi_{jt}^h + \psi_{it}^h \qquad \text{(d.2)}$$

where ψ_{ij}^h, ψ_{jt}^h and ψ_{it}^h are respectively the time invariant preferences of consumers in country *j* for the good *h* produced by country *i*, the time varying preferences of consumers in country *j* for product *h* and the time varying preferences for good *h* produced by *i*, across the consumers in all the countries considered. Because $z_{ijt}^h \phi$ are supposed to be greater or equal to zero, we assume the values of all parameters ψ^h being distributed as exponential.

We set δ_{jj}^2 unique across all the goods *h*, time *t* and country preferences *ij*. In order to have values of α_{ij}^h gathered in the positive domain we bound the value of δ_{jj}^2 to a set of maximum values[43] imposing an upper constraint to the prior distribution. From a theoretical perspective, we are limiting the irrationality and non-persistent component of national preferences. The distribution of σ_h and γ_h is a function of exponential, as in the previous setting. Values of β^h are assumed to be distributed as an exponential when we compute the non-normalized distance matrix W. When W is row-normalized[44], as usual in spatial econometric, β^h can be interpreted as the average influence of neighbors and their values are constrained in the (-1;+1) domain, to have positive semi-definite covariance matrix. β^h priors are therefore modelled as a Uniform distribution in the bounded space.

To estimate the parameters of our model we recur to variational inference (Blei et al., 2017) and specifically to the ADVI algorithm implemented by Kucukelbir et al. (2017). Variational inference turns the task of computing a posterior into an optimization problem finding the member of distributions that minimizes the Kullback-Leibler (KL) divergence to the exact posterior. ADVI transforms constrained latent variables to unconstrained real valued latent variables and computes derivatives of the joint distribution, expressing the gradient as an expectation over the family of the distributions and reparametrizing the gradient in term of a standard Gaussian. To compute Monte Carlo approximations we need therefore only to sample from a standard Gaussian. To simplify the simulation process, we estimate our parameters with a mean field Gaussian approximation that assume zero correlation among the transformed unconstrained latent variables[45].

Results for the non-normalized inverse distance matrix are shown in Table E.1. These are experimental results given that most of the estimations (especially when $\delta_{jj} > 0.001$) do not achieve stable convergences. In these cases, values around the minimum KL divergence are used.

---

[43] Preference parameters must be greater or equal to 0. Excluding zero flows trade, in order to avoid Inf and NaN values for the observed variable (given that we are considering the ratio of trade flows), restricts the domain of α_{ij}^h to strictly positive values. In our maximization process, we set α_{ij}^h equal to 1E-06 if the value sampled from the distribution is lower or equal to 0. Changing this lower bound does not affect significantly the result of our simulations. Imposing low constrains on the value of σ_{jj}^2 reduce frequency of negative values, increasing the stability of simulations.

[44] In a row-normalized distance matrix each element is equal to $\frac{(1/d_{jk})}{\sum_k (1/d_{jk})}$.

[45] As a robustness test, we compute the parameters using even a full rank Gaussian approximation, where the covariance matrix is estimated using a Cholesky factorization, to ensure the matrix remaining positive semi definite.

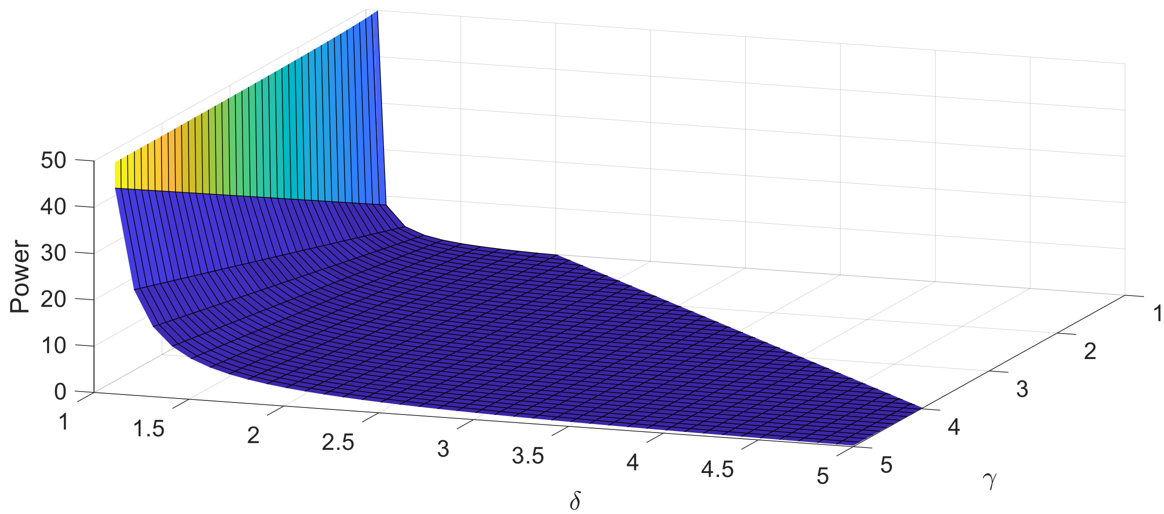
Table E.1 – Simulation of the correlation parameter β^h, residual variance δ_h and elasticity of substitution σ_h for sector 22, 87, 61, 38 using an inverse non-normalized distance matrix W.

Sector	δ_{jj} prior bounded to	Median	SD	2.5 Quartile [HPD 2.5]	97.5 Quartile [HPD 97.5]	δ_h: Median	σ_h: Median
β^{22}	0.1	6.332	5.583	1.086	20.512	2.743	4.682
β^{22}	0.01	16.319	90.31	5.27	39.36	1.785	3.109
β^{22}	0.001	12.651	5.247	5.329	25.135	0.969	2.284
β^{87}	0.1	3.706	1.988	1.193	8.697	2.477	4.396
β^{87}	0.01	7.153	1.988	4.002	11.894	1.899	4.417
β^{87}	0.001	20.909	6.793	10.49	36.764	0.72	2.282
β^{61}	0.1	5.166	3.708	1.116	15.175	2.343	4.862
β^{61}	0.01	26.515	13.213	9.497	60.112	1.212	2.563
β^{61}	0.001	21.317	9.955	8.071	46.22	0.613	1.871
β^{38}	0.1	5.854	4.728	1.131	18.071	2.662	4.527
β^{38}	0.01	14.732	7.961	4.963	34.469	1.913	3.578
β^{38}	0.001	17.496	7.943	6.709	37.086	1.048	2.007
β^{69}	0.1	5.363	3.947	1.127	15.64	2.546	4.675
β^{69}	0.01	11.924	6.118	4.176	27.626	1.846	3.522
β^{69}	0.001	6.548	2.584	2.886	12.857	1.271	2.443

Note: This table displays simulation of the spatial correlation parameter (β^h) specified in Eq. d.1. The residual variance (δ_h^2) is the variance of the error term for the distribution $N\sim(\mu, \delta_h^2)$ approximating our dependent variable $\ln\left(\frac{X_{ijt}^h}{X_{kjt}^h}\right)$, with μ specified in Equation 25. Lower value of $\delta_h$imply a better fitting of the distribution. Priors are $\beta^h \sim exp(0.2)$, $\delta_h \sim HalfCauchy(5)$, $\delta_{jj} \sim Unif[0, Bound(\delta_{jj})]$. W is the inverse distance matrix whose elements are $(1/d_{jk})$. Values are computed using variational inference (Blei et al., 2017), with ADVI algorithm (Kucukelbir et al., 2017).

With a large constrain on the variance of the consumers' preference shock (δ_{jj}^2), the unexplained variance (δ_h^2) surges. The intuition is that increasing the upper bounds of δ_{jj}, estimates for σ_h rises, as shown in Figure E.1, because of the structure of Equation 25, in an attempt to minimize the random component of preference.

Figure E.1 – Value of $\frac{\gamma_h}{(\sigma_h - 1)}$ conditioned to the joint distribution of $\gamma_h > \sigma_h > 1$



As a result, δ_h growths and $\frac{d\, ln\left(X_{ij}^h / X_{kj}^h\right)}{d\beta^h}$ decrease. Consequently, even the proposed value of β^h stuck near their initial values. Estimates for σ_h and, as a consequence of the nested structures, γ_h are indeed higher increasing the upper limit of δ_{jj}. Conversely, with low values for the upper limit of δ_{jj}, estimates of γ_h and σ_h tend to converge to the values previously estimated in Equation 30, as shown in table 5.

Table E.2 – Estimates of β^h for the subset of sectors with different parameter specification

Sector	$\boldsymbol{estimates\ of\ eq.3.30}$		$\boldsymbol{estimates\ of\ eq.3.30\ and\ of\ eq.3.d.1}$ with σ_{jj} bounded to 0.001 and W not normalized		$\boldsymbol{stimates\ of\ eq.3.30\ and\ of\ eq.3.d.1}$ with σ_{jj} bounded to 0.001 and W row-normalized	
	$\boldsymbol{\beta^h}$	σ_h	$\boldsymbol{\beta^h}$	σ_h	$\boldsymbol{\beta^h}$	σ_h
22	8.1	1.66	12.7	2.28	0.282	2.374
87	63.4	1.67	20.9	2.28	0.455	2.019
69	6.9	1.73	6.6	2.44	0.18	2.065
38	51.9	1.81	17.5	2.07	0.511	1.713
61	47.6	1.87	21.3	1.87	0.509	1.821

Note: Table D.2 show estimate of the spatial correlation parameter (β^h) of Equation 25 with preference parameters specified in Eq. d.1.