The Determinants of US Olive Oil Imports

A. Malek Hammami, John C. Beghin

Working Paper 22-WP 643
December 2022

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

A. Malek Hammami is Postdoctoral Research Associate, Gulf Coast Research and Education Center, University of Florida, Wimauma, Florida 33598. E-mail: ahammami@ufl.edu

John Beghin is CARD Faculty Fellow and Professor, Department of Agricultural Economics, University of Nebraska - Lincoln, Lincoln, Nebraska 68583. E-mail: beghin@unl.edu

For questions or comments about the contents of this paper, please contact John Beghin, beghin@unl.edu.

The determinants of U.S. olive oil imports

A. Malek Hammami and John C. Beghin*

This version: December 14, 2022

Abstract: We investigate the determinants of U.S. bilateral imports of olive oil and their dynamics from shocks in foreign supplies and changes in U.S. olive oil demand, using an augmented gravity framework that leads to an equilibrium of bilateral trade flows from olive oil exporters to the U.S. market. The empirical specification is applied at the disaggregated HS-6 level in a panel dataset, and three estimation techniques (truncated OLS, PPML, Heckman), for which the latter two account for zero trade flows, the extensive margin of trade and the potential censored distribution of exports with zero trade flows. We run Reset and HPC tests to qualify our results. On the supply side, exporters' capacity to exports, multilateral trade resistance, and immigrants' networks into the US are strong determinants of the bilateral trade flows for both aggregate olive oil exports and for virgin olive oil exports, On the consumer side, U.S. GDP, the import unit value, and immigrant network effects are robust determinants of bilateral flows as well for aggregate and virgin olive oil trade flows. Migrants' stock, exporters' GDP and population, and total exports revenues increase the probability of an exporter entering the U.S. market. We could not find robust evidence of consumer behavior being influenced by popular press measures of the emergence of Mediterranean diet and olive oil, or measures of cultural globalization of U.S. consumers.

JEL Codes: F14, Q17
Keywords: olive oil, trade, gravity equation, migrant network

*Hammami is affiliated with the Gulf Coast Research and Education Center, University of Florida, Wimauma, Florida. Beghin is affiliated with the Agricultural Economics Department and the C. Yeutter Institute of International Trade and Finance, at University of Nebraska Lincoln, and the Center for Agricultural and Rural Development at Iowa State University. Beghin is corresponding author: beghin@unl.edu. We thank Bo Xiong for discussions, Dinos Giannakas, and Dick Perrin for comments on an earlier draft, and participants at the 2021 AAEA meetings.

1. Introduction

Olive oil exports to the US have been increasing considerably for several decades. Trade flows have quadrupled in the last three decades (see table 1). Numerous factors might be at play to explain this strong growth of olive oil trade between the US and the rest of the world, from factors influencing import demand and export supply of olive oil. On the demand side, beyond simple demographic changes, income growth, and price effects, olive oil is known for its health benefits, and it is part of Mediterranean diet, which has been increasingly popular in the US. The presence of large immigrant populations of Mediterranean origin may also have helped popularizing the use of olive oil. The composition of olive oil imports has also evolved over time, towards higher quality imports of virgin olive oil and away from pomace, and more diversified sources. New export suppliers have emerged and entered the growing U.S. market. Network effects may have helped establishing olive oil business networks through Mediterranean migrants as it has been the case in other industries (Combes et al., 2005; Rauch, 1999). Hence, the extensive margin of trade (new products, new exporters) is another interesting aspect to investigate to explain the rapid evolution of these U.S. imports of olive oil.

We investigate U.S. imports of olive oil taking into account demand and supply shifters and elements of extensive margins using an augmented gravity-equation equilibrium framework. The framework incorporates usual demand shifters (prices, demographics, and income), the evolving sophistication of U.S. diet, bilateral and multilateral trade costs, and supply elements explaining the intensive and extensive margins of trade from various sourcing countries, into an equilibrium framework at the sectoral level (Yotov et al., 2016; and Anderson and van Wincoop, 2003). The framework leads to an empirical specification, which is then applied to a panel dataset at a disaggregated HS-4 and 6 levels for olive oil products, and with two estimation techniques, which account for the large number of zeros, the extensive margin of trade, and the potentially censored distribution of bilateral trade flows. We also run truncated OLS for comparison's sake.

Despite its importance, international trade of olive oil between the US and the rest of the world has received limited attention. Xiong et al. (2014) estimated U.S. demand for olive oil including the role of popular diet, distinguishing three olive oil types. Ronen (2017) investigated global aggregate olive oil trade looking at the impact of nontariff measures, using a gravity-like framework. Hammami and Beghin (2021) analyze the impact of retaliatory tariffs imposed by the US on olive oil imports sourced in the EU and Spain in particular.

Related to the supply side, we find that exporters' capacity to exports, multilateral trade resistance, and immigrants' networks from exporting countries into the US are strong determinants of the bilateral trade flows for both aggregate olive oil exports and for virgin olive oil exports. Relative to demand determinants, we find that U.S. GDP, the import unit value, and immigrant network effects are robust determinants of bilateral flows as well for aggregate and virgin olive oil trade flows. Regarding the extensive margin of trade, migrants' stock, exporters' GDP and population, and total exports revenues increase the probability of an exporter entering the U.S. market. Beyond the important result on migrant networks, we could not find robust evidence of systematic influences on U.S. consumer behavior by variables proxy-ing for the popularity of Mediterranean diet, or increasing popularity of olive oil, or measuring cultural globalization of U.S. consumers.

The following sections provide some background information on the olive oil sector in the US, and then describe the key elements of the conceptual equilibrium framework of the gravity equation with the relevant specifics of the investigation. Estimation techniques and data description follow, and findings are presented in the last section before conclusions.

2. Background on the U.S. olive oil market and olive oil exporters

2.1. The evolution of the U.S. olive oil market

Table 1 shows U.S. olive oil production supply and disappearance and documents the phenomenal growth of the market. Olive oil consumption has quadrupled since 1990 (USDA PS&D, 2020). As a result, the US has become the world first importer of olive oil (representing 10% of world production), with more than 90% of its domestic consumption being imported.

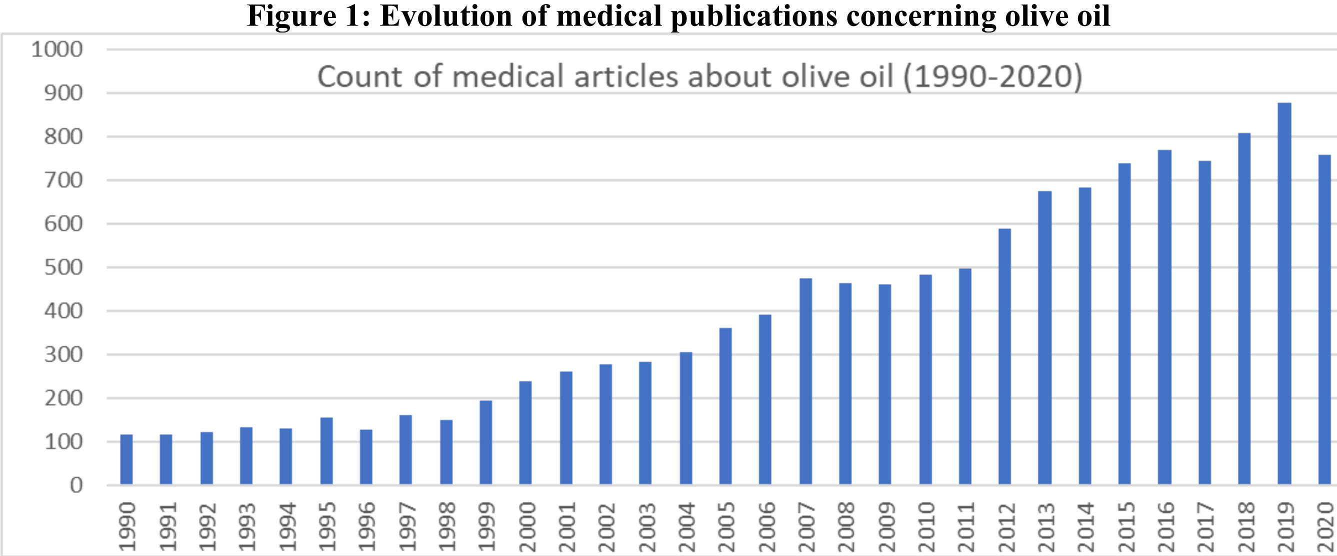
A series of cultural elements may have influenced the consumption of olive oil by U.S. consumers. First, interest in and knowledge about the benefits from olive oil, Mediterranean diet, and healthy diet have been continuously increasing among Americans (Pubmed.gov, 2020; Google Trends, 2021). Cultural globalization may also have facilitated the move away from Anglo-Saxon diet to a more Mediterranean one. Deeper influences may have come through cultural network effects with rising populations of immigrants from Mediterranean countries in which olive oil is paramount in the diet. They can influence US consumers' preferences and also facilitate business link back home to export to the US. Hence, we can hypothesize that migrant networks may have had influenced the adoption, level of consumption, and availability of olive oil in the US and its sourcing.

Table 1: U.S. olive oil Production, Supply and Disappearance in 1000 tons

Attribute	**1990/1991**	**1995/1996**	**2000/2001**	**2005/2006**	**2010/2011**	**2015/2016**	**2016/2017**	**2017/2018**	**2018/2019**	**2019/2020**	**2020/2021(P)**
Production	1	1	0	2	5	14	15	16	16	16	16
Imports	100	114	212	242	290	330	316	322	355	390	380
Total Supply	101	115	212	244	295	344	331	338	371	406	396
Exports	4	11	4	9	4	8	13	12	7	6	10
Consumption	97	104	208	235	291	336	318	326	364	400	386
Distribution	101	115	212	244	295	344	331	338	371	406	396

To capture the growing stock of health knowledge on olive oil, we rely on Pubmed.gov to compute an index of the number of published academic refereed articles in medical journals mentioning key search terms (olive oil). This index allows for a longer and less biased series than those based on internet data.[1] As shown in Figure 1, the index shows a pattern of growing number of articles mentioning olive oil (Pubmed.gov, 2020). We also conjecture that the popularity of olive oil could come from cultural influence of migrants from olive-oil producing countries. We show the stock of migrants in Figure 2, both for the simple total and for olive-oil import-weighted stock. The stock of migrants suggests a strong correlation with the increasing consumption of olive oil. The bilateral nature of the migrant panel data provides more variation than the number-of-articles variable which only varies over time.

Figure 1: Evolution of medical publications concerning olive oil

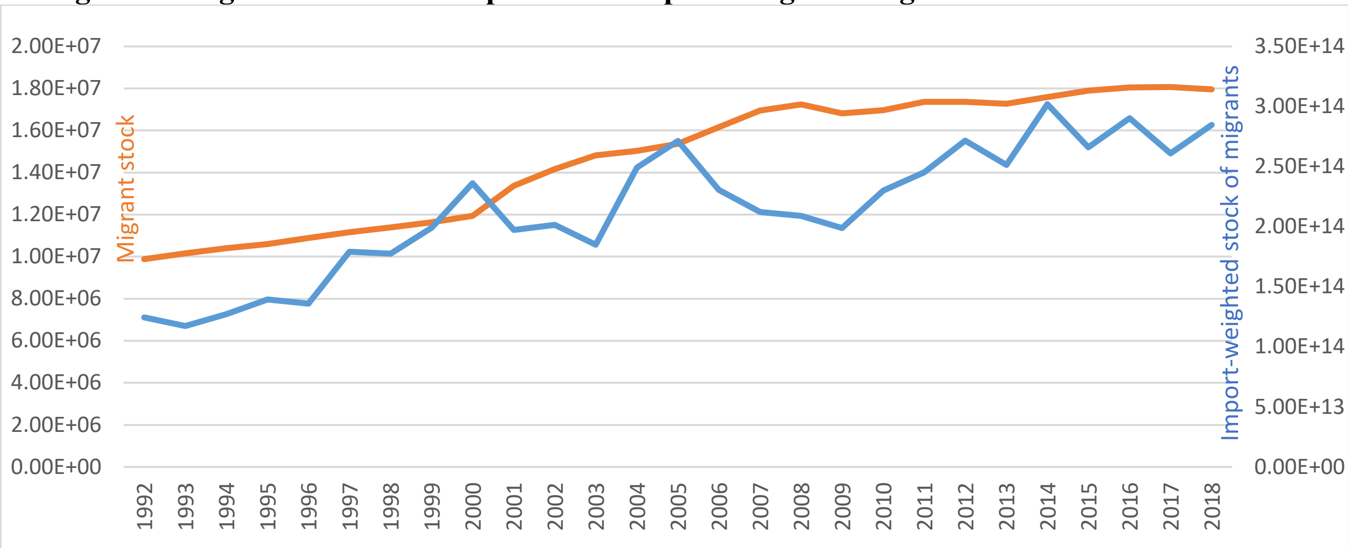


2.2. Patterns of U.S. imports of olive oil

The olive oil market in the US can be differentiated into two main categories: virgin and non-virgin oils. Virgin oil is considered a higher quality product, since during this process olives have been simply pressed with no heat or chemicals involved.

[1] In the econometric estimation, we also use a related index reflecting the stock of popular press articles on Mediterranean diet using https://www.newsbank.com/ as in Xiong et al.

Figure 2: Migrant stock from exporters & import-weighted migrant stock in the US



The oil is pure and not refined. All other olive oils, heat or chemically treated, are considered non-virgin and can be sometimes mixed with some virgin oil and simply called olive oil. Olive oil extracted by chemical process is called pomace and is the lowest quality product. Virgin olive oil itself has many subcategories: Extra virgin (cold press), first cold pressed, and organic. (IOC, 2020; and Vossen, 2007).

The "Olive Oil & Its Fractions" category imported to the USA under the HS code 1509, has the largest average share of consumption of more than 90% of all olive oils imported and consumed in the US. The remaining share is the "Olive-residue Oil & Blends" category under HS code 1510 (edible and non-edible). The "Olive Oil & Its Fractions" category (HS code 1509) divides into "Virgin olive oil/fractions" category (HS code 150910) and "Refined olive oil / fractions" (HS code 150990) category. The virgin olive oil (HS 150910) includes extra-virgin, labelled and organic fractions which are of a superior quality than the refined one. Since the early 1990's, the share of virgin olive oil has been increasing from 35% to reach almost 80% of the olive oil imports. This increase reflects both the rising consumption of olive oil and the shift towards higher quality olive oil consumed in the US. The global economic crisis of 2008-09 temporally reset the clock on this evolution as shown in figures 3a and 3b; trends are clear.

Figure 3.a: Evolution of import shares: Olive oil and Residual (1992-2019)

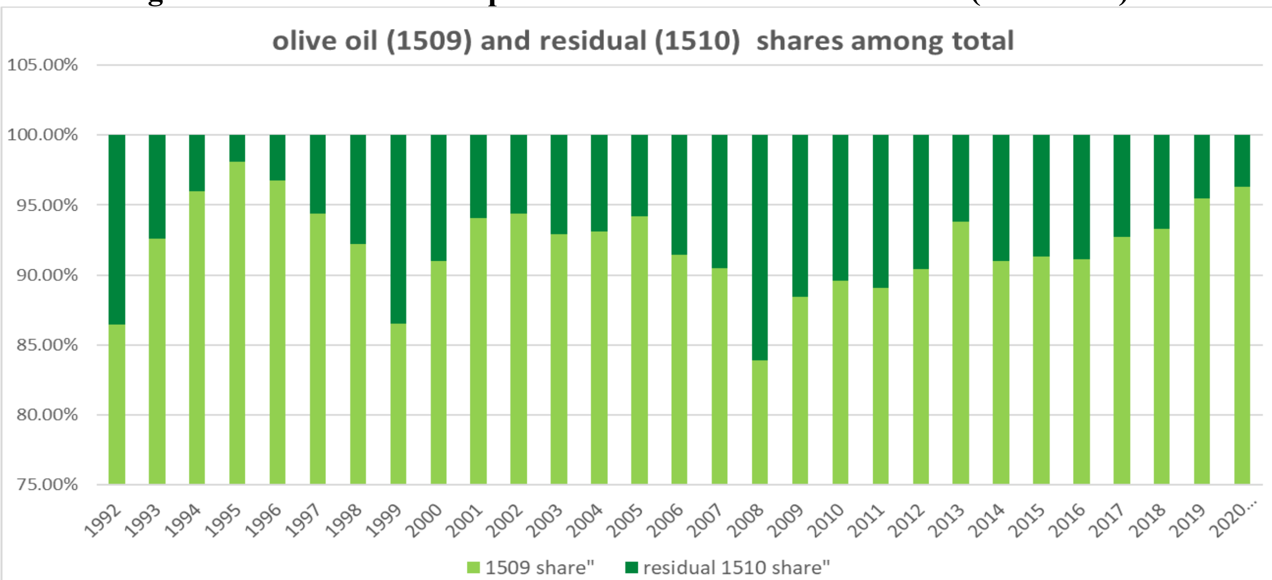


Figure 3.b: Evolution of import shares: Virgin and Refined (1992-2020)

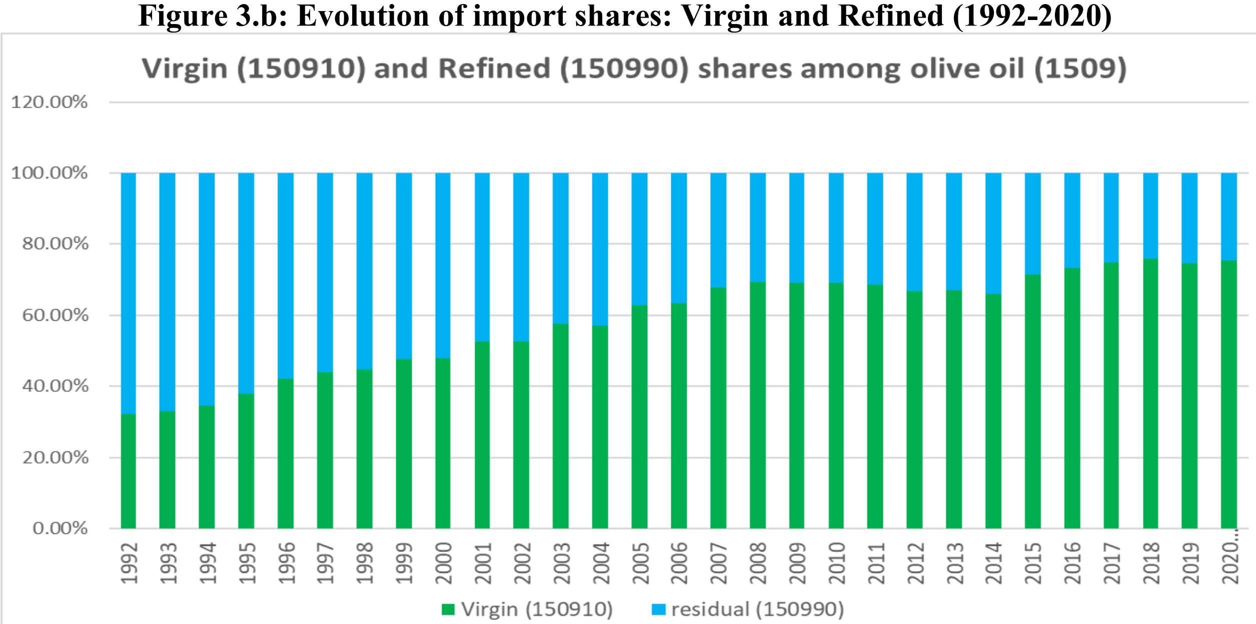
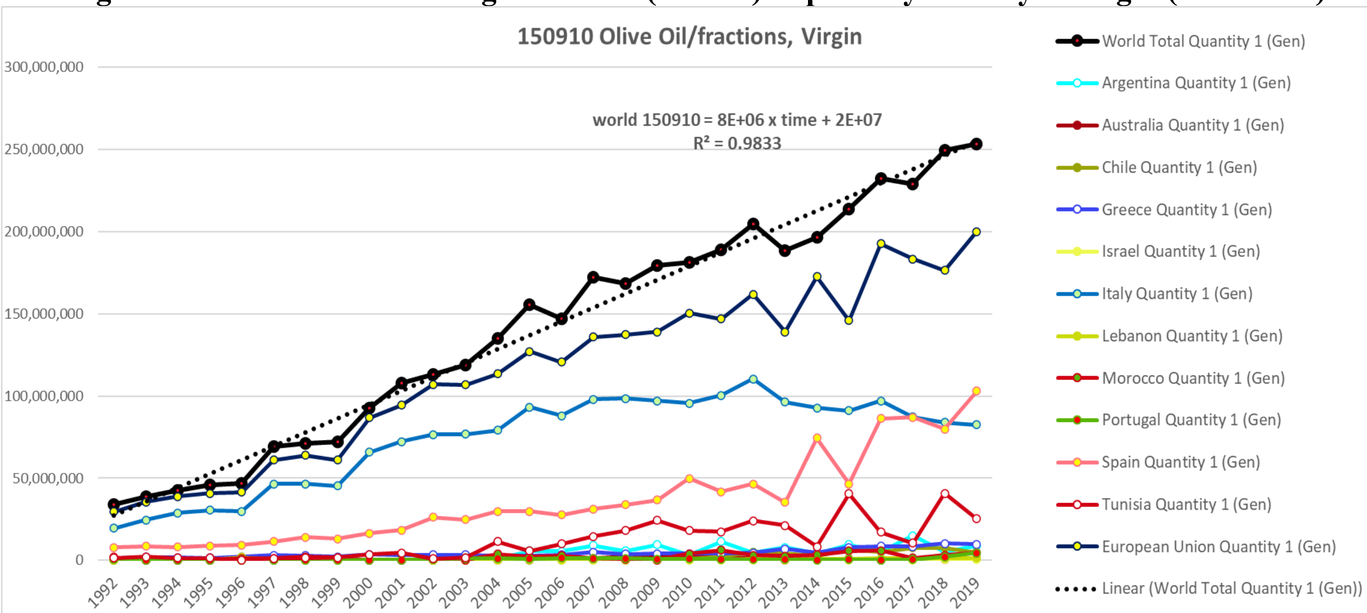


Figure 4 shows the evolution through time of U.S. imports for virgin olive oil by import source. EU sources dominate (Spain and Italy), but the rising importance of non-EU Mediterranean sources (Tunisia, Morocco) and the existence of a competitive fringe (Argentina, Israel, Lebanon, Chile, Australia) are also noted.

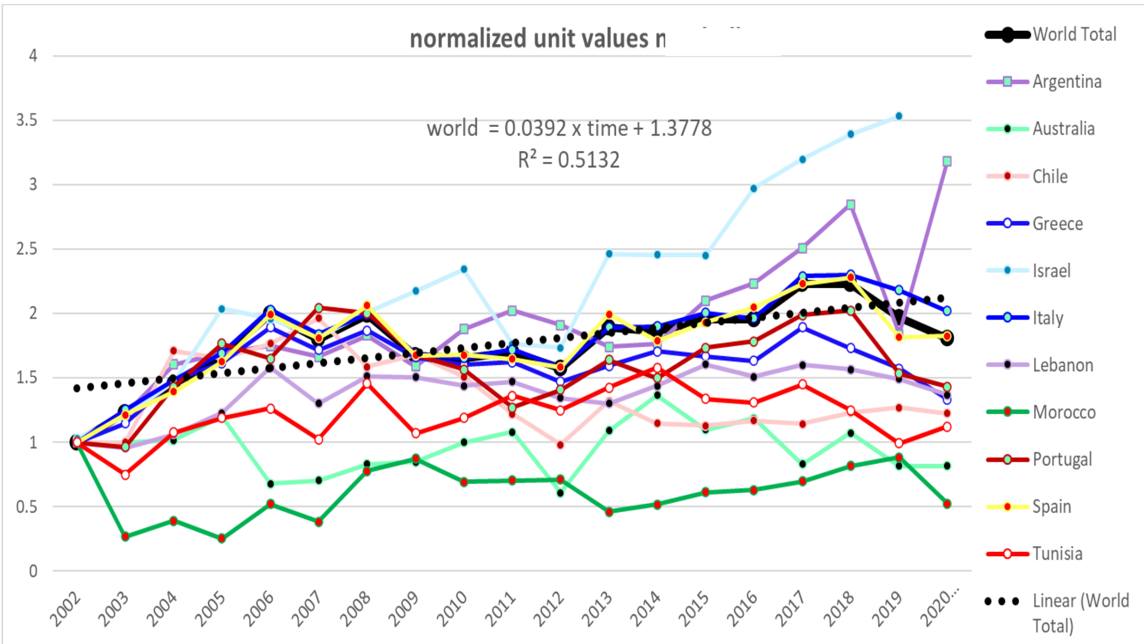
Figure 4: Evolution of U.S. virgin olive oil (150910) imports by country of origin (1992-2019)



Most of the virgin olive oil imported to the US is from Italy and Spain. Spain overtook Italy in 2018. The rest of the countries exporting to the US have remained small exporters but with growing quantities exported. As a group they provide a competitive fringe to the established exporters. Tunisia has increased its exports to the US the most, since 2004, approaching Spain exports in 2015, and remains the largest exporter within the fringe of smaller exporters. Argentina has had a noticeable increase since 2005 and is now the 4th largest virgin olive oil exporter to the US, after Italy, Spain, and Tunisia. Olive production for olive oil exhibits stochastic yields with "good and bad" years resulting in annual variations of production even for established exporters. Inventories partially mitigate these variations. Variations in export supply are also reflected by significant variations in import unit values, as shown in Figure 5. The Figure shows normalized nominal import unit values to a common base (2002=100).

In aggregate, the average nominal import unit value has been increasing over time by roughly 4% per year. The increase is small, and but slightly above inflation. The 2008-9 crisis had an impact. The relative normalized prices show increasing disparities among exporters. The dispersion of import unit values reflects different qualities and differentiation among competitors over time. The empirical investigation leverages this variation in import unit values. Italian, Israeli, and Argentinian unit values have been increasing faster than the other ones.

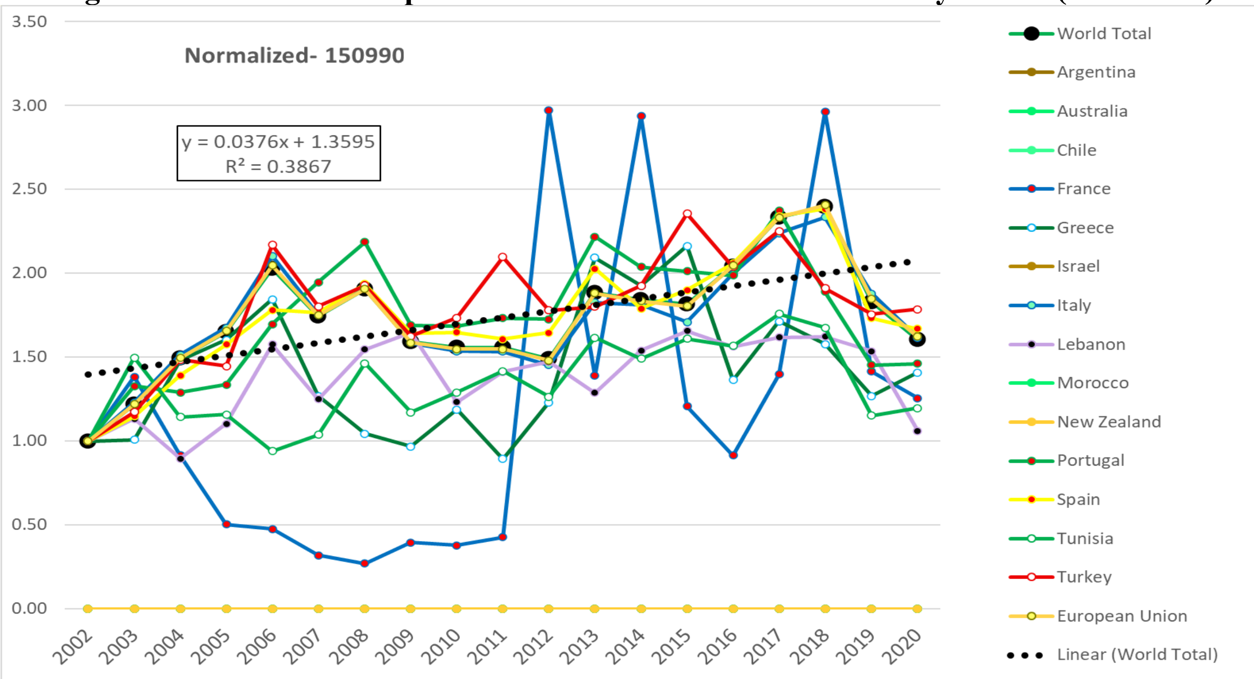
Figure 5: Virgin import Normalized unit value evolution in U.S. imports by country (2002-2019)



U.S. imports of refined olive oil have been on a decreasing trend for most exporters (not shown) as consumers upgraded to virgin olive oil. We also note that leading exporting countries for virgin olive oil are the ones leading exporters of refined olive oil to the US. Italy and Spain are leaders ahead of the others (fringe), with Spain overtaking Italy since 2013. The fringe of other exporters still has Tunisia as the third largest source competing with Turkey and Morocco in a close level.

Figure 6 presents normalized nominal import unit values to a common base (2002=100) for refined oil imports. The world unit value tends to increase between 2002 and 2020 with an average 3.8% rate of change. Dispersion across sources is smaller than for the normalized virgin olive oil unit values.

Figure 6: Refined U.S. import Normalized unit value evolution by source (2002-2019)



2.3.Supply shocks in World olive oil markets

Various supply shocks and changes interact with U.S. demand of olive oil. Producing countries, endowed with a specific Mediterranean climate, are competing among themselves to supply the world market, including the US. Profit-maximizing firms in these countries' supply chains compete and adapt to changing market conditions. New entries and production techniques have put pressure on average unit values. Spain, as an example, invested hugely on reforming olive oil production and opted for an intensive production since the 1960's. Nowadays, Spain supplies almost half of world production (46%) (Guerrero, 2014).

Italy and Greece relied both on their historical reputations and their authentic ancient know-how to signal their quality. Italy imports large amounts of Spanish oil, which find their way back on the world market. With the development of digital marketing strategies and globalization, there are evidence that olive oil producers are going up market and that non-traditional producers are entering the international market. As a result, there is a glut in world's supply of olive oil which pushes producers to differentiate their product for a higher quality (IOC (2020); Vossen (2007); Lavee (2007); Milli (2006); USDA PS&D (2020)). The increasing and now dominant share of virgin olive oil in U.S. imports reflect this fact.

Finally, olive oil production has a stochastic yield due to environmental and agronomic shocks. Weather, pathology, and physiological state of the trees impact its yearly production. Olive trees are biennial trees that have alternate yearly production. One year above the average production and one year below (Lavee, 2007). We later investigate this potential variability of yield, although unit values of traded olive oil reflect that variability to a great extent.

2.4. Evidence on the evolution and sophistication of consumer demand

Consumers around the world including in the US, have been increasingly concerned about the quality of food they purchase (IFIC, 2018). Many studies have investigated the relation between health and nutrition information and food demand. Early research about health and nutrition factors has found evidence of diversion of U.S. demand from food containing cholesterol and heavy fats. Early on, Brown & Schrader (1990) showed how cholesterol information has affected shell egg consumption using both fixed and changing coefficients models. Capps & Schmitz (1991) implemented a Rotterdam demand model that accounts for health and nutrition factors in meats consumption; Chern et al. (1995) used a Bayesian model of health risk belief and consumer awareness surveys on healthy primary food to reveal saturated fats information's effects on oils and fats consumption. Piggott & Marsh (2004) investigated consumer response to publicized food safety information on the U.S. meat demand using a Generalized Almost Ideal Demand Model.

Other studies have approached health information and demand from the experimental and behavioral perspective. Park & Davis (2001) provided theoretical reasons under which Instrumental variable approaches may not be superior to OLS when analyzing cross-sectional health information, when instruments are poor and measurement error is inconsequential. Hilger et al. (2011) used a retail field experiment to highlight the importance of experts' opinion on experience goods demand, like wine. There has been a general agreement that food nutrients and health information have been influencing food demand. Through years, food consumers have been increasingly interested in healthy food, thus, increasing the quality and sophistication of their food consumption.

Many relatively new diet trends have been emerging with various news coverage and success, such as keto, vegan, vegetarian, and Mediterranean diet, among others. The latter has been considered by the popular press as the best overall diet for the last two consecutive 2019 and 2020 (CNN, 2020). Mediterranean diet has spread in the industrialized world from its origin in the Mediterranean basin. Alexandratos (2006) and Regmi et al. (2004) investigated this rise of the Mediterranean diet and related it to globalization and income growth. The former used data from FAO's food balance sheets to assess the historical evolution of the Mediterranean diet. The latter examined previous literature and trends of global and U.S. food consumption determinants are examined. Regmi et al. (2004) analyzed trade data to determine changing diets phenomenon's effect on Mediterranean diet products' trade.

Studies of olive oil demand in North America have been scarce (Del Giudice et al., 2015). Xiong et al. (2014) estimated U.S. demand for olive-oil differentiated products using the AIDS model and accounting for the impact of information on Mediterranean diet. They find that both the stock and number of press articles discussing olive oil and health is positively related to the level of olive oil imports. They aggregate all olive oils imports into three aggregate categories. They do not account for possible influence through migrant networks. Menapace et al. (2009) study olive demand in Canada through a survey that demonstrated the significance of geographical indication and certification of origin Label.

Main studies on olive oil focus on Europe. Most of them are surveys and experiments (Karipidis et al., 2005; Kalogeras et al., 2009; Bernabéu & M. Díaz, 2016; Cacchiarelli et al., 2016; Carbone et al., 2018; and Scarpa & Del Giudice, 2004). Del Giudice et al. (2015) summarized the literature about consumer preferences for extra-virgin olive oil attributes in a scoping review. They used a 'narrative systematic review' of the topic and then derived a willingness to pay for key attributes inspired by the literature's review.

Relevant to our analysis, in the context of a net-importing country (with small domestic production), Kavallari et al. (2011) investigate the structure of import demand of Germany and UK for olive oil from southern European producers. That study itself has been based on Vlontzos and Duquenne (2008) on Greek olive oil potential in the international market. Finally, Garcia Álvarez-Coque and Martí Selva (2006) using a gravity model to estimate euro-Mediterranean fruits and vegetable trade flows.

3. Model

3.1. The gravity equation approach to bilateral trade

Anderson (1979) developed the theoretical foundations of the gravity model, which was then elaborated by Eaton and Kartoum (2002); and Anderson and van Wincoop (AvW) (2003), among others. AvW (2003) provide a simple model based on supply endowment and representative CES preferences, symmetric trade costs and multilateral trade resistance terms reflecting trade costs faced by exporters and importers in a cross-section data context. The approach is widely popular among trade economists, despite its drawbacks being overly structured with symmetric trade costs, fixed endowment approach to supply and its normalization of price limited to cross-section data (Baldwin and Taglioni, 2006). These assumptions are relaxed here. The AvW approach when used in general equilibrium tends to overestimate trade costs as shown by Balistreri and Hillberry (2007).

Methods of estimation have evolved greatly to address the presence of zero bilateral trade flow, the extensive margin to trade, and typical estimation mistakes biasing results. Baldwin and Talglioni (2006) identified three principal mistakes of gravity model's applications in the literature, the so-called golden, silver and bronze medal mistakes. The Golden medal mistake occurs when multilateral trade resistance terms are omitted and are correlated with error terms, leading to bias. The bronze medal mistake occurs when bilateral trade flow volumes are wrongly deflated by a common deflator. The silver medal mistake is the wrong log-averaging of the bilateral trade flow volumes. The authors proposed several time-varying and unvarying dummies as an attempt to adjust gravity regression mistakes.

More complex issues with gravity models such as the extensive margin of trade have been addressed as well. Several authors have formalized the extensive margin of trade (Melitz (2003); Chaney (2008) and Helpman et al. (2007)). The central idea uses heterogeneous firms and the most productive firms enter new markets at the extensive margin. Yotov et al. (2016) develop a supply-based gravity model allowing for more substitution in production with a CET production frontier differentiated by destination. The model leads to an isomorphic specification of the gravity similar to the CES-demand based approach, unless specific demand and supply shifters are available to differentiate the CET and CES elasticities (Xiong and Beghin, 2014; Cadot et al., 2018). Gould (1994), Rauch (1999), and Combes et al. (2005) focus on network effects on trade for differentiated products. These dimensions are present in our approach.

The basic gravity model

As in AvW, we assume homothetic preferences (we only have a single destination j (*j*=US)). Products are differentiated by country of origin (Armington, 1969). The demand in country j, is obtained from maximizing a CES-utility function, with utility derived from consuming products differentiated by origin i $(all\ exporters\ of\ olive\ oil)$. The setup extends to a sectoral approach from which we abstract here to simplify the presentation. The maximized utility is:

$$\left\{\sum_i \ a_i^{\frac{1-\sigma}{\sigma}} c_{ij}^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{\sigma}{\sigma-1}},$$

Subject to the following budget constraint for a set expenditure E

$$\sum_i p_{ij} c_{ij} = E_j \ ,$$

where, $\sigma > 1$ is the elasticity of substitution between goods. The exogenous taste parameter a_i is the CES preference parameter, which will be instrumental later to incorporate the impact of diet information. The consumption of varieties from country i in country j is given by c_{ij}. Total expenditure in country j (E_j) measured at delivered prices $(p_{ij} = p_i \ t_{ij})$defined as a function of factory-gate prices in country of origin$(p_i)$and bilateral trade costs markup (1+trade cost) from i to j $(t_{ij} > 1)$.

Consumers maximize their utility from consuming countries i's goods under budget constraint. This yields:

$$x_{ij} = \left(\frac{a_i p_i t_{ij}}{P_j}\right)^{(1-\sigma)} E_j \ , \qquad (1)$$

where $x_{ij}$are trade flows in value from origin i to destination j, and P_j denotes a CES consumer price index or "inward multilateral resistance" as defined by AvW (2003), which illustrates ease of market access into j:

$$P_j = \left[\sum_i \left(a_i p_i t_{ij}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \qquad (2)$$

Finally, the derivation imposes a market clearing condition for each exported good for which the production is equal to the sum of demand including domestic demand. The equilibrium condition assumes that the shipped quantities "melt" on the way to their destinations by an amount equivalent to the trade cost. This yields the following:

$$Y_i = \sum_j x_{ij} = \sum_j \left(\frac{a_i p_i t_{ij}}{P_j}\right)^{1-\sigma} E_j, \qquad (3)$$

where, $Y_i = p_i Q_i$ is the nominal value of production in country i. It is equal to factory-gate prices p_i multiplied by Q_i supply of the given product in exporter i before melting of the iceberg occurs (representing trade cost). Bilateral trade cost is an iceberg cost. For the consumers to receive x_{ij}, exporters have to send $x_{ij}\, t_{ij}$ with $t_{ij} > 1$. Trade flows melts in the way by $(t_{ij} - 1)$ and buyers bear that extra-cost of $x_{ij}\,(t_{ij} - 1)$. For every exporting country and every product, we assume that production of country i is $Y_i = p_i Q_i = \sum_j x_{ij}$. Equation (3) presents the determinants of the equilibrium trade flow between countries i and j. The term $(a_i p_i)^{1-\sigma} = \frac{Y_i}{\Pi_i}$, is used to eliminate the factory price in the equilibrium condition. The price index Π_i is given by:

$$\Pi_i \equiv \sum_j \left(t_{ij}/P_j\right)^{1-\sigma} E_j\,, \qquad (4)$$

which is the "outward multilateral resistance" that shows exporter i's ease of market access into all j countries (AvW, 2003, Baldwin and Taglioni, 2006). Substituting (3) and (4) the structural gravity system is presented as:

$$x_{ij} = \frac{Y_i E_j}{\Pi_i}\left(\frac{t_{ij}}{P_j}\right)^{1-\sigma}, \qquad (5)$$

where $Y_i E_j$ illustrates the "size term", and $\frac{1}{\Pi_i}\left(\frac{t_{ij}}{P_j}\right)^{1-\sigma}$ is the "trade cost term". If $t_{ij} = 1$ for all i

and j, then $P = \Pi$ and trade is maximized since outward and inward trade resistance are equal with minimum trade costs. In our sectoral application, we assume that Y_i represents the capacity to export olive oil by country i. Then for any exporter i the ratio $\frac{1}{\Pi_i}\left(\frac{t_{ij}}{P_j}\right)^{1-\sigma}$ represents the relative trade cost of the U.S. market relative to all destination markets served by that exporter. We note that the ratio is varying monotonically with the ratio of the U.S. import unit value for that exporter and the average real import unit value of that exporter to all destinations.[2] We use this element in our empirical strategy.

3.2. Empirical strategy (estimation approach and empirical specification)

As shown in section 2, the set of countries exporting olive oil to the US has been changing over time. Some countries made their market entrance several years ago, such as Australia, Brazil, Algeria, Peru, Slovenia, among others. When a competitor has not yet entered or chooses to exit the U.S. market for a given year, its trade volume will be taking the value of zero. In general, to accommodate zeroes, many investigations use Poisson pseudo maximum likelihood (PPML), which allows to include zeroes as part of the intensive margin of trade by taking the exponent of equation (5) and the logarithm of continuous variables on the right-hand-side of the equation.

One can add an extensive margin to this which is confounded with the intensive margin and does not address potential selection into exporting to a destination market and the potential censoring in zero observations. PPML estimates are consistent under heteroskedasticity and the approach provides a natural solution to mechanically handle zero values of the dependent variables and provides a robust covariance matrix estimator (Santos Silva and Tenreyro, 2006).

[2] One can multiply and divide the ratio by the ex factory price p_i raised to the power of (1-σ) to obtain $\frac{1}{\Pi_i}\left(\frac{p_i\, t_{ij}}{p_i\, P_j}\right)^{1-\sigma} = \frac{1}{\dot{\pi}_i}\left(\frac{p_i\, t_{ij}}{P_j}\right)^{1-\sigma}$, with $\dot{\pi}_i \equiv \sum_j \left(p_i t_{ij}/P_j\right)^{1-\sigma} E_j$. The latter expression is a transform of the weighted average import unit value for exporter i to all its destinations j, with weights being the income of the importers rather than their imports. Under homothetic and identical preferences the income and import levels are proportional.

We address the extensive margin of trade using Heckman sample selection as in Helpman et al. (2008) and Santos Silva et al. (2015), which accounts for the censoring of the latent variable representing the decision to trade or not through its first-stage Probit. Helpman et al. (2008) have a second correction for the extensive margin of heterogeneous firms, which we do not have here. For completeness we also provide truncated OLS estimates, using the strictly positive observations in the dataset. We also run a Probit of the probability to export by countries over time, to gauge the fit of the selection equation in the Heckman model. We use Stata.

For Statistical tests, first, a Ramsey RESET test is applied to check if the models exhibit evidence of misspecification. Second, we use the HPC test proposed by Santos Silva et al., (2015) to choose which of PPML or Heckman fits our olive oil data the best. As detailed in the appendix, the HPC test allows discrimination between a pair of competing models to fit data with many zeros (Santos Silva et al., 2015). The test allows to compare "two-process models" like Heckman to a simpler 1-process model accommodating zeros without explicit extensive margin like the PPML approach.

Specification and variable proxies

Since we have panel data (1992-2018) for 21 exporters, we add a time subscript to our specification. Each olive oil k imported by the US at year t from exporter i, has the following trade flow equation: $X_{iUStk} = \frac{Y_{it}E_{US,t}}{\Pi_{ikt}}\left(\frac{t_{iUStk}}{P_{USkt}}\right)^{1-\sigma}$. (6)

We also make use of the ratio indicated in footnote 1. Since both the U.S. import unit value and the exporter's average import unit value are time varying, our combined variable $\frac{1}{\pi_{it}}\left(\frac{p_{it}\, t_{ijt}}{P_{jt}}\right)^{1-\sigma}$ is obviously time varying. For product k, it is proxy-ed by the ratio $\frac{UV_{i,US,t,k}}{avgUV_{i,t,k}}$ where $avgUV_{i,t,k}$ is the average unit value of product k over all destinations served by exporter i. It plays the role of outward trade resistance measure since it includes all the t_{ij} trade costs for all j markets served by exporter i.

Further, we use the time-varying capacity to export olive oil of exporters to approximate output y_{it}; $Tot\_exports_{i,t,k}$ is exporter i's total exports of olive oil k to the world in year t. It is a solid proxy for exporter i's production and capacity to export olive oil k. We choose this variable as olive production data from FAO are incomplete for several countries and do not disaggregate olive oil types. The smoothing role of inventories is also complicate and implicit in total exports. We also address potential endogeneity of supply determinants using a series of exporter fixed effects to address some of the potential endogeneity issues coming from omitted variables. This is not implementable on the importer side since we only have a single destination in the trade flow data (j=USA).

On the demand side of equation (6), we already have the import unit value $UV_{i,t,k}$ and then we use 2010-price constant U.S. GDP, $realGDP_{US,t}$ as a proxy for consumers' income. Note that the latter income measure includes demographic change affecting U.S. GDP.

Regarding trade-flow shifters related to diet adoption, cultural influences and market penetration, we first rely on immigration networks (stock of migrants from olive oil exporting countries). Immigration, and tourism are important factors having network effects that can affect food trade flows (Kavallari et al. (2011)). Rauch and Trindade (2002) found that the share of ethnic Chinese populations as immigrants affects bilateral trade of that country with China. This is consistent with earlier investigations (Gould (1994) and Rauch (1999)). Rauch (1999) used proximity variables such as distances –this was prior to the CEPII database availability on distance. We investigate this geographic proximity relying on the geodist database of CEPII. Following Gould (1994) and Combes et al. (2005), we look at migrant networks coming to the

US. The immigration data are collected from the Office of Immigration Statistics' (OIS) yearbooks of immigration statistics 1996-2019. We use tables of Persons Obtaining Lawful Permanent Resident Status by Region and Country of Birth (Office of Immigration Statistics, 2021). In the Heckman model with the extensive margin, we further investigate the role of migrants into the extensive margin. The latter would capture business network influences, rather than their cultural influence, of migrant networks from olive-oil producing countries.

We rely on established strategies to incorporate the impact of health information and information on popular diets in demand systems. Variables capturing U.S. consumer's demand sophistication are incorporated into a CES framework through preference parameter a_i in equation (1). Indices reflecting sophistication of demand are as follows. First, we use the KOF index of cultural globalization which measures the degree of globalization of 122 countries over time based on economic, social and political criteria on a 100-score scale for each country. The index was first introduced by Dreher (2006) at the Konjunkturforschungsstelle at ETH Zurich, Switzerland from which it takes the name "KOF" (kof.ethz.ch, 2020; Gygli et al. 2019). We use the index value for the US which varies over time. Next, we use a PubMed index which is based on counting the number of published scientific refereed articles about the searched terms (olive oil, healthy diets, etc.). This index only counts the health and medical publications. We use both annual flow and accumulated stock of publications (see pubmed.gov, 2020). We also develop a popularity index reflecting the stock of news articles on Mediterranean diet and health benefits associated with olive oil based on the NewsBank database as in Xiong et al. (2014).

In some additional runs reported in the appendix, we investigate adding yield to capture technical change elements in olive oil production not captured by the unit value. Yield data is incomplete and leads to a decrease in observations.

We look at the impact of regional trade agreements on olive oil trade. Tariffs on olive oil are typically low and do not vary much over time, especially in the context of a single destination. However, deeper market access may have an impact on trade flows via trade facilitation measures. We also look at lead and lag effects of RTAs on olive oil trade flows. These results are put in appendix as they did not exhibit any robustness once exporter fixed effects are introduced (see appendix table A.1). The limited number of RTAs for the US may also explain the lack of significant results.

Trade flow data are based on UN COMTRADE. Observations were restricted to the 21 top exporters (Algeria, Argentina, Australia, Brazil, Chile, Croatia, Cyprus, Egypt, France, Greece, Israel, Italy, Jordan, Lebanon, Morocco, Portugal, Slovenia, Spain, Tunisia, Turkey, and Peru) representing in excess of 99.8% (for 2019) of imports to the US. Excluded countries from the initial selection are: Albania, Canada, China, Japan, Mexico, Montenegro, New Zealand, Palestine, South Africa, and Syria. These countries were excluded because several data series were missing and also because of market disruptions, such as in Syria. The panel dataset is available for total exports, imported quantities and prices (import unit values). It extends from 1992 to 2018 consisting of olive oil (HS:1509) and its two major components: virgin oil (HS: 150910) and refined oil (HS:150910). Pomace or residual olive oil (HS:1510) is dropped from the estimation because of too many missing observations and also because of its limited use in food consumption. We use physical quantity trade data (variable c_{iUStk} rather than value X_{iUStk}) since we have both physical unit data and import values in COMTRADE. To sum up, our panel consists of 26 years (1992-2018) and 21 countries, which results in 546 total observations. Depending on the olive oil type, there is a considerable number of zeros and some missing values that could be dropped during regressions.

We have the final empirical specification of trade flows as follows:

$$c_{iUStk} = f\left(realGDP_{US,t}, \frac{UV_{i,US,t,k}}{avgUV_{i,t,k}}, Tot_{exports_{itk}}, fixed\ effect_i, migrant\ stock_{it}, taste\ shifter_{ust}\right) (7).$$

Equation (7) is run in the typical exponential form of the log of continuous variables and fixed effects for the PPML estimation. For the Heckman and truncated OLS specifications, a double log specification of (7) is run. For the Heckman specification, the inverse Mill's Ratio coming from the covariance between selection and level equations is also included as an explanatory variable. We specify the selection equation as follows.

$$L_{iUStk} = f\left(\begin{matrix} realGDP_{US,t}, \frac{UV_{,i,US,t,k}}{avgUV_{i,t,k}}, Tot_{exports_{itk}}, fixed\ effect_i, migrant\ stock_{it}, taste\ shifter_{ust}, \\ pop_{i,t}, GDP_{i,t}, Cons_{US,t}, Exprev_{i,t}, EU_{i,t} \end{matrix}\right) (8),$$

where, L_{iUStk} is the trade latent variable. It takes the following values: $\begin{cases} L_{iUStk} = 1 \text{ when } c_{iUStk} > 0 \\ L_{iUStk} = 0 \text{ when } c_{iUStk} = 0 \end{cases}$.

Variable $pop_{i,t}$ is the exporter's population, $GDP_{i,t}$ is the exporter's GDP, $Exprev_{i,t}$ is the exporter's total olive oil export revenues, and $EU_{i,t}$ is a dummy that indicates whether the exporter is a European Union member, all expressed at a time period t. We also included the U.S. consumption of olive oil at a time t, $Cons_{US,t}$, to capture the growth of the market over time. Note that the part-one Probit (or the selection equation) has the same variables as the main Heckman (part-two) equation, with additional variables explaining the decision to export or not and which are excluded from the trade level equation for proper identification.

Dealing with missing observations

For a number of observations, the import unit values $UV_{i,US,t,k}$ (or in a few cases $avgUV_{i,t,k}$)

values were missing which would have resulted in omitting these observations from regressions. This is the case for instance for the zero trade flow observations which do not have an observed unit value to the U.S. market. We use two instruments to replace missing values of $UV_{i,US,t,k}$, which represent plausible expectations of potential exporters in country i. First, we use the average import unit value of the two geographically closest countries in the sample that are exporting to the US for that year. Distance data from CEPII allow us to compare country pairs distance-wise. For example, Cyprus has the following countries from the sample ordered by increasing distance: Lebanon, Israel, Jordan, Egypt, Turkey, and Greece. We compute its missing unit values as follows: $UV_{CYP,US,t} = Average(UV_{LBN,US,t}, UV_{ISR,US,t})$. If Lebanon and Israel also have missing unit values, recovering them would include Cyprus again. For that reason, two of them need to have, near complete set of unit values or must depend on another group of countries farther than the two closest with existing trade flows. Second, we assume that the missing UV to the US is expected to be equal to the average unit value for that country and year ($UV_{i,US,t}$ $=avgUV_{i,t}$), which means the "expected" import unit value of i olive oil to the US equals the average unit value for its worldwide shipments for that year. The choice of instrument has virtually no impact on the results.

4. Results

Table 2 presents the estimation results for the three chosen estimations methods (Truncated OLS, PPML and Heckman sample selection (with its Probit selection equation)). All specifications include exporter fixed effects, which are not reported to save space. Country fixed effects are significant, except for France, Croatia, Jordan, and Portugal. Peru is omitted in the intercept. Results are available upon request. The three oil categories are run as separate regressions, although in separate runs with the disaggregated categories (HS 150910 and 150990) we included cross-price effects which were not significant.

The three models have a strong R^2 for olive oil HS:1509 and its subcategories (>0.9). The unit value ratio as a price proxy for trade cost and exporter trade resistance has the expected negative sign, except for the refined lower quality oils HS:150990 that has an insignificant positive sign. Similarly, the GDP as proxy for income has a systematically significant and positive sign, which confirms the aggregate income effect for olive oil products in the US. The refined olive oil HS:150990 with its lower quality has an insignificant income coefficient.

Among non-price shifters, the candidate variables were tried with mixed success. First, network effects appear to be very robust in all equations, the stock of migrants from exporting countries living in the US explain much of the variation of olive oil imports over time and across export sources. For the shift of taste parameters, we tried several combinations of the KOF index and PubMed and NewsBank indices, both in stock and flow form. However, we could not find robust results, the estimations were plagued by sign reversals, and loss of significance. In addition the inclusion of these shifter in the specification tends to dilute the significance of the income variable. Hence, we are not confident that we can capture the alleged taste changes with these indices, beyond the cultural influence captured by the migrant network variable. Detailed results are reported in appendix. Table 2 below illustrates the type of insignificant results which we obtained (here with the KOF index).

A comparative horizontal reading of the Table 2 permits a cross quality comparison. The higher quality virgin olive oil (HS:150910) has generally a higher price and income response than its lower quality analogue (HS:150990). Migrants' stock has almost the same effect on both subcategories. Olive oil (HS:1509) that regroups both subcategories, shows effects that are close to the average of the effects on two subcomponents at the HS-6 level, which is expected and reassuring.

The Heckman sample selection specification is also shown in Table 2. Its first-stage Probit explains the extensive margin for exporters who decide to export or not. The Probit has an acceptable R^2 of around 0.6 in most runs. It shows that migrants' stock, exporters' GDP, exporters' population, and total exports revenue are among significant factors that influence decisions to initiate exports of olive oil to the US. This result also provides some insight on the business-network of migrant to create new trade (the extensive margin) as opposed to their influence on the intensive margin through their expanding cultural influence. Heckman's second stage (trade levels) provides results comparable to PPML and truncated OLS results. Coefficients' signs and robustness are similar to PPML, with stronger price and income response for the higher quality virgin olive oil HS:150910. Heckman's Mills' ratio significance suggests that sample selection is present among the observations.

Table 3 presents Reset tests for the three models shown in Table 2. For the aggregate olive oil (HS:1509), the RESET tests suggest that the truncated OLS is fine as well as the Heckman trade volume equation. The Reset test for PPML is borderline significant at 10% but not at 5% or less. We read it as inconclusive. For the disaggregated regressions (150910 and 150990), the RESET tests indicate that truncated OLS, PPML, and Probit have evidence of misspecification. The Heckman trade volume equation fails the reset test for 150910 but passes the test for 150990. Table 4 presents the HPC tests in a comparison between PPML and Heckman sample selection models to discriminate the best model (one-part versus two- part models) for our data. The HPC test uses predicted trade levels but based on a logarithmic estimation for the Heckman and in trade levels for PPML. The Heckman estimated values are then reconverted in levels for the test, following Santos Silva et al. (2015).

Table 2: Comparative Gravity estimations OLS-PPML-PROBIT & HECKMAN

	1509				*150910*				*150990*			
	OLS	*PPML*	*PROBIT*	*HECKMAN*	*OLS*	*PPML*	*PROBIT*	*HECKMAN*	*OLS*	*PPML*	*PROBIT*	*HECKMAN*
ln_gdp_USA	1.010**	0.828***	-1.273	1.093***	2.151***	2.350***	0.231	2.299***	-0.166	0.931**	-0.029	0.397
	(0.012)	(0.004)	(0.709)	(0.007)	(0.000)	(0.000)	(0.949)	(0.000)	(0.796)	(0.023)	(0.992)	(0.710)
ln_IUV	-1.310***	-1.052***	-0.037	-1.337***	-1.168***	-1.295***	-0.158	-1.224***	-0.968***	0.516*	0.039	-0.924***
	(0.000)	(0.000)	(0.841)	(0.000)	(0.000)	(0.000)	(0.546)	(0.000)	(0.000)	(0.052)	(0.825)	(0.000)
ln_Exporter's total_Exports	0.762***	0.826***	-0.512	0.824***	0.600***	0.722***	0.014	0.682***	0.254***	0.379**	0.237	0.414***
	(0.000)	(0.000)	(0.104)	(0.000)	(0.000)	(0.000)	(0.963)	(0.000)	(0.000)	(0.042)	(0.238)	(0.000)
ln_Stock_mig	0.828***	1.145***	0.761***	1.029***	0.413*	0.973***	0.685***	0.732***	1.385***	1.341***	0.325***	1.764***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.086)	(0.000)	(0.000)	(0.008)	(0.000)	(0.000)	(0.003)	(0.001)
ln_KOFCuGIdf	0.0508	0.382	-2.492	-0.113	-0.991	-0.308	0.195	-1.116	0.969	-1.560***	-0.51	0.291
	(0.946)	(0.304)	(0.269)	(0.880)	(0.239)	(0.221)	(0.926)	(0.227)	(0.431)	(0.003)	(0.753)	(0.883)
ln_pop_Exporter			-0.92***				-0.97***				-0.631**	
			(0.002)				(0.004)				(0.012)	
ln_gdp_Exporter			0.736**				1.090**				0.748***	
			(0.033)				(0.011)				(0.003)	
ln_US_consumption			1.323				0.545				0.655	
			(0.403)				(0.737)				(0.629)	
ln_tot_exp_rev			0.900***				0.261				-0.049	
			(0.006)				(0.395)				(0.808)	
member_eu_Exporter			0.020				0.023				0.244	
			(0.972)				(0.966)				(0.460)	
Exporters' Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mill's ratio				0.767***				1.008***				1.930**
				(0.004)				(0.000)				(0.022)
_cons	-37.95***	-40.69***	12.907	-45.84***	-62.04***	-80.26***	-45.724	-72.43***	-9.214	-33.81***	-23.25	-30.92
	(0.000)	(0.000)	(0.887)	(0.004)	(0.000)	(0.000)	(0.637)	(0.000)	(0.561)	(0.005)	(0.768)	(0.258)
N	449	542	534	534	413	515	512	512	332	484	456	480

R-sq (pseudo)	0.937	0.981	0.610	-	0.921	0.984	0.632	-	0.874	0.956	0.498	-

P-values in parenthesis. * p=<0.1; ** p<0.05; *** p<0.01. Note that country specific fixed effects of this regression were deleted from this table for simplicity reasons.

Table 3: RESET test

	1509				*150910*				*150990*			
	OLS	PPML	PROBIT	HECKMAN	OLS	PPML	PROBIT	HECKMAN	OLS	PPML	PROBIT	HECKMAN
RESET test P-Value	0.792	0.0684	0.203	0.652	0.000	0.0017	0.0238	0.0000	0.0000	0.0000	0.0237	0.8333
Chi2/F-value	0.35	3.32	1.62	0.20	8.22	9.81	5.11	18.71	9.71	32.95	5.12	0.04
N	449	542	534	534	413	515	512	512	332	484	456	480
R-sq (pseudo)	0.937	0.981	0.61	-	0.921	0.984	0.63	-	0.874	0.956	0.498	-

Table 4: HPC test

	1509		*150910*		*150990*	
HPC test	*PPML*	*HECKMAN*	*PPML*	*HECKMAN*	*PPML*	*HECKMAN*
P-Value	0.000	0.133	0.000	0.181	0.000	0.535
T-value	9.280	1.114	14.694	0.913	9.663	-0.087
N	534	534	511	512	477	480
R-sq (pseudo)	0.981	-	0.984	-	0.956	-

The test clearly rejects PPML, while not rejecting Heckman for the aggregate trade and disaggregated flows as well. We conclude that both RESET and HPC tests suggest the Heckman sample-selection model provides the best fit for our data to address the U.S. import demand determinants for olive oil, compared to PPML. The Inverse Mills Ratio is also significant which suggests that the error terms of the selection equation and the trade equation are correlated. The extensive margin contribute to explaining why exporting countries enter or not the U.S. market for olive oil. Regardless of the method of estimation, results are consistent across these approaches. The selection equation informs uniquely on the role of migrant network on the extensive margin of trade.

Beyond the specifications presented in Table 2, multiple variations in specifications were tried. We added yield, to capture technical change. However, several countries had missing yield data, leading to a decrease in observations. Results were not robust with sign reversal and various level of significance. Regional trade agreements (RTA) between the US and its partners were also introduced, including with lag and lead effects to attempt to capture potential effects of deeper trade integration. (See appendix table A.1 for an example of RTA specification). Results again were not robust. Both were removed from our preferred specification because exporters' fixed effects capture much of these two variables, except their time variation.

In addition, bilateral distance which, is supposed to be one of the main proxies for trade cost, was tried and found not to be robust. It exhibited sign reversal (significant positive sign), the inverse of what the theory suggests. We conjecture that the lack of variation in destinations combined with the geography of olive production led to this result. Most of olive oil exporters are located in the Mediterranean basin (say as opposed to Mexico and Canada for other commodities) and quality differs from an exporter to another with new-world exporters (Chile, Argentina) are just emerging relative to further-distant established EU exporters. Our unit cost ratio variable incorporates the cost of shipping the various olive oils and in that sense is a solid proxy for several trade costs including the cost of overcoming distance.

Finally, for both subcategories of olive oil (virgin and refined), we have thought of testing cross-price effects in the specifications. It appeared that cross-prices were not significant with no effect on the explanatory variables nor the R^2.

5. conclusion

This paper investigated the determinants of U.S. import demand and export supply of olive oil in an equilibrium framework. We used an augmented gravity equation equilibrium framework. On the demand side, income and relative prices are important determinants of olive oil demand. Migrant networks exhibit a systematic influence on demand via a cultural element in the intensive margin of trade reflecting new preferences for olive oil. These migrant communities and networks propagate new culinary habits and food consumption, in this case the Mediterranean diet and olive oil. Beyond this strong result related to migrant networks, and despite using several alternative proxies to measure the influence of information and cultural influences on U.S. olive oil demand, we could not find any additional robust association, somewhat to our surprise. This result is in contrast to Xiong et al. We use a much finer disaggregation of olive oil by origin (21 exporters) relative to Xiong et al. (3), who abstracted from the influence of migrant network in their context. The strong aggregation in Xiong et al. smooths out much of the variability we have in the bilateral flows and allows to identify a trend, probably too difficult to identify in our bilateral flows.

On the supply side, exporters' specific fixed effects were used to absorb the cross-sectional variation among olive oil exporters to the U.S. market. The variability of export supply over time and geography was captured by the time-varying value of exports to all destinations for each exporter. In addition, the relative trade costs of destination markets was captured by the ratio of the import unit value to the US relative to those of other destinations, and this for each exporter.

We also found strong network effects at the extensive margin, created by immigrant from exporting countries. These networks facilitate new trade flows, beyond their influence on demand via the intensive margin in business networks lowering the cost of entry into the US market. Other variables influencing the extensive margin include the exporter's GDP, population, and total exports revenues, as suggested by the Probit results.

From an empirical point of view, the Heckman sample selection specification was the best model to fit the data and did not exhibit evidence of misspecification as suggested by the RESET test. The HPC test suggests that the Heckman sample selection with its explicit extensive margin fits the data better than the PPML approach does. PPML provides a legitimate way to incorporate zero but lacks the extensive margin component.

Olive oil (HS:1509) and its subcategories: virgin (HS:150910) and refined (HS:150990) were the olive oil types analyzed in the investigation. In future investigations and using shorter time series, we plan to further disaggregate olive oil at the HS 8 or HS 10 levels. In addition, a decomposition of bulk/non-bulk packaging, and a consideration of organic and labelled subcategories of olive oil could help provide deeper insights on the quality upgrade which took place over time in the U.S. olive oil market.

References

Alexandratos, N. (2006). The Mediterranean diet in a world context. *Public Health Nutrition*, *9*(1a), 111-117.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American Economic Review*, 69(1), 106-116.

Anderson, J. E., & Van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, *93*(1), 170-192.

Armington, P. S. (1969). A theory of demand for products distinguished by place of production.

Staff Papers, 16(1), 159-178.

Baldwin, R., & Taglioni, D. (2006). Gravity for dummies and dummies for gravity equations (No. w12516). *National bureau of economic research*.

Balistreri, E. J., & Hillberry, R. H. (2007). Structural estimation and the border puzzle. *Journal of International Economics*, 72(2), 451-463.

Beghin, John C., and Heidi Schweizer. (2021) "Agricultural Trade Costs." *Applied Economic Perspectives and Policy* 43, no. 2s: 500-530.

Bernabéu, R., and M. Díaz. (2016). "Preference for olive oil consumption in the Spanish local market," *Spanish Journal of Agricultural Research* 14(4): e0108.

Brown, D. J., & Schrader, L. F. (1990). Cholesterol information and shell egg consumption. *American Journal of Agricultural Economics*, *72*(3), 548-555.

Cacchiarelli, L., A. Carbone, T. Laureti, and A.Sorrentino. (2016). "The Value of different Quality Clues in the Italian Olive Oil Market," *Italian Review of Agricultural Economics* 71(1): 372-379.

Cadot, O., J. Gourdon and F. van Tongeren. (2018). "Estimating Ad Valorem Equivalents of Non-Tariff Measures: Combining Price-Based and Quantity-Based Approaches," *OECD Trade Policy Papers*, No. 215, OECD Publishing, Paris. http://dx.doi.org/10.1787/f3cd5bdc-en.

Capps Jr, O., & Schmitz, J. D. (1991). A recognition of health and nutrition factors in food demand analysis. *Western Journal of Agricultural Economics*, 21-35.

Carbone, A., L. Cacchiarelli, and V. Sabbatini. (2018). "Exploring quality and its value in the Italian olive oil market: a panel data analysis," *Agricultural and Food Economics* 6(1): 6.

Chaney, T. (2008). Distorted gravity: the intensive and extensive margins of international trade. *American Economic Review*, 98(4), 1707-21.

Chern, W. S., Loehman, E. T., & Yen, S. T. (1995). Information, health risk beliefs, and the demand for fats and oils. *The Review of Economics and Statistics*, 555-564.

CNN (2020). Best and worst diets for 2020, ranked by experts, with a popular one near last https://www.cnn.com/2020/01/02/health/best-diet-worst-diet-2020-wellness/index.html . Retrieved 10/15/2020.

Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer. (2005). "The trade-creating effects of business and social networks: evidence from France." *Journal of international Economics* 66.1: 1-29.

Del Giudice, T., Cavallo, C., Caracciolo, F., & Cicia, G. (2015). What attributes of extra virgin olive oil are really important for consumers: a meta-analysis of consumers' stated preferences. *Agricultural and Food Economics*, 3(1), 20.

Del Giudice, T., Cavallo, C., Caracciolo, F., & Cicia, G. (2015). What attributes of extra virgin olive oil are really important for consumers: a meta-analysis of consumers' stated preferences. Agricultural and Food Economics, 3(1), 1-15.

Dreher, Axel (2006): Does Globalization Affect Growth? Evidence from a new Index of Globalization, *Applied Economics* 38, 10: 1091-1110.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5), 1741-1779.

García Alvarez-Coque, J. M. and Martí Selva, M. (2006), "A Gravity Approach to Assess the Effects of Association Agreements on Euromediterranean Trade of Fruits and Vegetables", Working Paper 06/15, *TRADEAG*.

Google trends (2020) https://trends.google.com/trends/explore?date=all&geo=US&q=olive%20oil,%2Fm%2F02c7v1,%2Fg%2F1q6j9fx41 retrieved 10/02/2020.

Gould, David M. (1994). "Immigrant links to the home country: empirical implications for US bilateral trade flows." *The Review of Economics and Statistics* (1994): 302-316.

Guerrero Marta (2014). Olive Oil Production in Spain Set to Rebound. *USDA/FAS. GAIN Report,* number: SP1402. https://apps.fas.usda.gov . Retrieved 10-14-2020.

Gygli, Savina, Florian Haelg, Niklas Potrafke and Jan-Egbert Sturm (2019): The KOF Globalisation Index – Revisited, *Review of International Organizations*, 14(3), 543-574. https://doi.org/10.1007/s11558-019-09344-2.

Hammami, A. Malek, and John C. Beghin. (2021). "The Trade and Welfare Impacts of the US Retaliatory Tariff on EU Olive Oil." Forthcoming in *Agricultural Economics*.

Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics* 123(2), 441-487.

Hilger, J., Rafert, G., & Villas-Boas, S. (2011). Expert opinion and the demand for experience goods: an experimental approach in the retail wine market. *Review of Economics and Statistics 93*(4), 1289-1296.

IFIC (2018). *2018 Food and Health Survey*. https://foodinsight.org/2018-food-and-health-survey/. Retrieved 09-27-2020.

International Olive Council (IOC) 2020. Designations and definitions of olive oils. https://www.internationaloliveoil.org/olive-world/olive-oil/ . Retrieved 10/02/2020.

Kalogeras, N., S. Valchovska, G. Baourakis, and P. Kalaitzis. (2009). "Dutch consumers' willingness to pay for organic olive oil." *Journal of International Food & Agribusiness Marketing* 21(4): 286-311.

Karipidis, P. I., E. Tsakiridou, N.M. Tabakis, and K. Mattas. (2005). "Hedonic analysis of retail egg prices." *Journal of Food Distribution Research*, vol. 36, no 856-2016-56442: 68-73.

Kavallari, A., S. Maas, and P. M. Schmitz. (2011). "Examining the determinants of olive oil demand in nonproducing countries: Evidence from Germany and the UK." *Journal of Food Products Marketing* 17, no. 2-3 (2011): 355-372.

kof.ethz.ch (2020). *KOF Globalisation Index*. https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html . Retrieved 10-04-2020.

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

Lavee, S. (2007, January). Biennial bearing in olive (Olea europaea L.). In *Annales Ser His Nat* Vol. 17, pp. 101-112.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6), 1695-1725.

Menapace, L., Colson, G., Grebitus, C., & Facendola, M. (2011). Consumers' preferences for geographical origin labels: evidence from the Canadian olive oil market. *European Review of Agricultural Economics* 38(2): 193-212.

Milli, S. (2006), "Olive Oil Marketing on Non-Traditional Markets: Prospects and Strategies", *New Medit* Vol. 5 (1): 27-37.

National library of medicine. Pubmed.gov (2020) https://pubmed.ncbi.nlm.nih.gov/ retrieved 10/02/2020.

Office of Immigration Statistics (2021), *Yearbook of Immigration Statistics*. https://www.dhs.gov/immigration-statistics/yearbook , retrieved 04/04/2021.

Park, J., & Davis, G. C. (2001). The theory and econometrics of health information in cross-sectional nutrient demand analysis. *American Journal of Agricultural Economics* 83(4): 840-851.

Piggott, N. E., & Marsh, T. L. (2004). Does food safety information impact US meat demand? *American Journal of Agricultural Economics 86*(1): 154-174.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society*: Series B (Methodological) 31(2): 350-371.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics* 48(1): 7-35.

Rauch, J. E., & Trindade, V. (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics* 84(1): 116-130.

Regmi, A., Ballenger, N., & Putnam, J. (2004). Globalisation and income growth promote the Mediterranean diet. *Public Health Nutrition*, *7*(7): 977-983.

Ronen, E. (2017). The Trade-Enhancing Effects of Non-Tariff Measures on Virgin Olive Oil. *International Journal of Food and Agricultural Economics 5*(3): 9-26.

Santos Silva, J. M. C., Tenreyro, S., & Windmeijer, F. (2015). Testing competing models for non-negative data with many zeros. *Journal of Econometric Methods* 4(1): 29-46.

Santos Silva, J.M.C., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4): 641-658.

Scarpa, R., and T. Del Giudice. (2004). "Market Segmentation via Mixed Logit: Extra-Virgin Olive Oil in Urban Italy," *Journal of agricultural & food industrial organization* v.2 no.1: pp. 20.

Vance, C., & Ritter, N. (2014). Is peace a missing value or a zero? On selection models in political science. *Journal of Peace Research* 51(4): 528-540.

Vlontzos, G. N., & Duquenne, M. N. (2008). Greek Olive Oil: How Can Its International Market Potential Be Realized? *Estey Journal of International Law and Trade Policy* 9(1753-2016-141197): 32-47.

Vossen, P. (2007). Current opportunities in the California olive oil industry. In California Plant and Soil Conference.

Xiong, B., and J.C. Beghin. (2014) “Disentangling the Demand-enhancing & Trade-cost Effects of Maximum Residue Regulations,” *Economic Inquiry*, 52 (3) (2014): 1190–1203.

Yotov, Y. V., Piermartini, R., Monteiro, J. A., & Larch, M. (2016). *An advanced guide to trade policy analysis: The structural gravity model*. Geneva: World Trade Organization.

APPENDICES (Not intended for publication)

HPC test

Davidson and MacKinnon (1981) developed a testing procedure for "corner solutions data" models with "zeros" or missing values of the dependent variable. The HPC test allows discrimination between two competing models to fit data with many zeros (Santos Silva et al., 2015).

Suppose a model A: $M_A: E[y_i|x_i] = g_A(x_i'\beta_A)\ F_A(x_i'\gamma_A)$,

To be tested against a model B: $M_B: E[y_i|x_i] = g_B(x_i'\beta_B)\ F_B(x_i'\gamma_B)$,

HPC uses a nested model C: $M_C: E[y_i|x_i] = (1-\alpha)g_A(x_i'\beta_A)\ F_A(x_i'\gamma_A) + \alpha g_B(x_i'\beta_B)\ F_B(x_i'\gamma_B)$,

where for $j = \{A, B\}$, $g_j(.) > 0, 0 < F_j(.) \leq 1$, vector parameters $\beta_j and\ \gamma_j$.

It tests $H_0: \alpha = 0$ and the alternative $H_0: \alpha = 1$, and permits three possible outcomes: accept both models, reject both models, and accepting one and rejecting the other. The test is structure such that it does not matter which model is chosen as A or B and vice versa.

Yield and RTA results (See table A.1)

Adding yield and RTAs would capture technical change and trade agreements between the US and its exporters. Total exports are included in the equation as a proxy for exporter's i capacity to export to the world in general. That gives an idea on its outward trade resistance of each exporter. Table A.1 shows the effects of adding yield and RTA to the basic equation (without fixed effects), since the Fixed effects would capture all the supply shocks, including yield, and bilateral shocks with the exporters, including RTAs. Their effect on trade flow is significant and positive for olive oil (HS-1509) and its both sub-categories (HS-150910 and 150990). That happens without big changes in basic equation's estimates, robustness nor its R2. We can also not that RTA agreements have a stronger effect than the yield does.

Taste proxies (see table A.2)

Several combinations of taste parameters have been tested to increase the explanatory power of the model. The best combination was “exporters’ migrants’ stock” and the cultural globalization index “KOFCuGIdf”. The most robust is the migrants’ stock which encompasses the network effect that increased import demand and spread Mediterranean cuisine. Indeed, immigration from exporters’ countries did increase their respective countries exports to the US. The cultural globalization index isn’t robust for the equilibrium equation (7) but showed some robustness in a simple demand function as in equation (4). Including other combinations of taste proxies like: healthy diet, Mediterranean diet or olive oil PubMed publications, had impact on the estimated effects of the unit value ratio and income effect without adding to the model’s explanatory power.

Table A.1: Basic, yield and RTA estimations

HS	*1509*					*150910*					*150990*				
	basic	yield	RTA	FE_yield	FE_RTA	basic	yield	RTA	FE_yield	FE_RTA	basic	yield	RTA	FE_yield	FE_RTA
Ln$GDP_{US,t}$	1.114***	1.059***	1.015***	0.856***	0.797***	2.572***	2.578***	2.527***	2.173***	2.153***	-0.447	-0.602	-0.505	-0.319	-0.299
	(-2.6)	(-2.45)	(-2.32)	(3.96)	(3.57)	(-7.56)	(-7.44)	(-7.23)	(11.16)	(9.69)	(-0.96)	(-1.33)	(-1.07)	(-1.39)	(-1.17)
Ln$\frac{IUV}{avgIUV}$	-3.011***	-2.851***	-3.142***	-1.713***	-1.700***	-4.288***	-4.182***	-4.339***	-1.750***	-1.826***	-0.751*	-0.644	-0.792	0.247	0.224
	(-13.11)	(-12.33)	(-13.37)	(-7.67)	(-7.34)	(-13.78)	(-12.89)	(-14.30)	(-6.98)	(-6.44)	(-1.62)	(-1.49)	(-1.59)	(0.88)	(0.75)
Ln $Tot_{exports_{itk}}$	0.955***	0.922***	1.003***	0.806***	0.839***	0.941***	0.908***	0.963***	0.676***	0.695***	0.941***	0.891***	0.973***	0.487**	0.556**
	(-26.8)	(-27.94)	(-24.15)	(12.10)	(13.15)	(-28.88)	(-30.05)	(-24.13)	(13.45)	(12.99)	(-16.6)	(-17.71)	(-15.21)	(2.65)	(2.76)
ln_Yied_o		0.208**		0.166*			0.162*		0.274***			0.412***		0.472***	
		(-2.27)		(1.80)			(-1.82)		(3.56)			(-2.76)		(3.28)	
agree_pta			0.903***		0.639***			0.328*		0.111			0.878***		1.047***
			(-4.7)		(2.61)			(-1.64)		(0.51)			(-3.12)		(3.29)
_cons	-34.82***	-34.11***	-32.80**	-26.04***	-23.87***	-78.93***	-79.71***	-77.98***	-65.29***	-62.70***	13.41	15.87	14.59	9.237	11.26
	(-2.70)	(-2.65)	(-2.50)	(-4.03)	(-3.66)	(-7.70)	(-7.75)	(-7.47)	(-11.51)	(-9.81)	-0.95	-1.17	-1.02	(1.37)	(1.52)
Exporter Fixed Effects	No	No	No	Yes	Yes	No	No	No	Yes	Yes	No	No	No	Yes	Yes
N	542	515	540	515	540	515	491	513	491	513	484	476	482	476	482
pseudo-R-sq	0.871	0.87	0.874	0.974	0.975	0.897	0.9	0.898	0.980	0.979	0.797	0.802	0.8	0.950	0.948

t-statistics in parenthesis. * p=<0.1; ** p<0.05; *** p<0.01

Table A.2: Taste proxies estimates

	1509					*150910*					*150990*				
	All proxies	Healthy	Med_diet	Oliveoil	KOFCuGIdf	All proxies	Healthy	Med_diet	Oliveoil	KOFCuGIdf	All proxies	Healthy	Med_diet	Oliveoil	KOFCuGIdf
$Ln\ GDP_{US,t}$	-1.39*	0.12	0.39	-0.035	0.83***	1.19*	1.97***	1.36**	0.98	2.35***	-1.55	0.27	-0.35	-0.82	0.93**
	(-0.06)	(-0.89)	(-0.63)	(-0.96)	(0.00)	(-0.08)	(0.00)	(-0.04)	(-0.18)	(0.00)	(-0.28)	(-0.82)	(-0.74)	(-0.50)	(0.02)
$Ln\ \frac{IUV}{avgIUV}$	-0.98***	-1.08***	-1.05***	-1.04***	-1.05***	-1.08***	-1.28***	-1.27***	-1.25***	-1.29***	0.64**	0.48*	0.48*	0.49*	0.52*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(-0.02)	(-0.08)	(-0.08)	(-0.07)	(0.05)
$Ln\ Tot_{exports_i}$	0.81***	0.82***	0.81***	0.81***	0.83***	0.69***	0.72***	0.72***	0.72***	0.72***	0.34*	0.41**	0.41**	0.41**	0.38**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(-0.06)	(-0.04)	(-0.04)	(-0.04)	(0.04)
$Ln\ mig\ stock_{it}$	1.23***	1.14***	1.16***	1.17***	1.15***	1.12***	0.97***	0.98***	0.99***	0.97***	1.62***	1.29***	1.29***	1.31***	1.34***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ln_KOFCuGIdf	0.13				0.38	-0.37				-0.31	-1.71**				-1.56***
	(-0.71)				(0.304)	(-0.155)				(0.221)	(-0.001)				(0.003)
ln_PubMed_HealthyDiet_stock	0.25	0.14				-0.55**	0.03				-0.74	0.01			
	(-0.41)	(-0.33)				(-0.05)	(-0.71)				(-0.22)	(-0.99)			
ln_PubMed_MedDiet_stock	-1.02**		0.07			-0.31		0.10			-0.33		0.07		
	(-0.02)		(-0.48)			(-0.52)		(-0.19)			(-0.62)		(-0.54)		
ln_PubMed	1.60***			0.17		1.25**			0.20*		1.71*			0.19	

_OO_stock															
	(-0.01)			(-0.27)		(-0.04)			(-0.09)		(-0.08)			(-0.36)	
Exporter's Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
_cons	18.84	-18.68	-26.42	-14.41	-40.69***	-49.30***	-70.46***	-52.59***	-42.22**	-80.26***	34.12	-20.38	-2.166	10.93	-33.81**
	(-0.37)	(-0.47)	(-0.28)	(-0.58)	(0.00)	(0.00)	(0.00)	(0.00)	(-0.04)	(0.00)	(-0.39)	(-0.55)	(-0.95)	(-0.76)	(0.04)
N	542	542	542	542	542	515	515	515	515	515	484	484	484	484	484
R-sq															
pseudo-R-sq	0.982	0.981	0.981	0.981	0.981	0.985	0.984	0.984	0.984	0.984	0.957	0.954	0.954	0.954	0.956
RESET															
chi2(1)	4.94	2.54	2.09	2.32	0.0684	9.71	16.08	16.52	15.65	0.0017	33.21	32.18	32.29	32.63	36.1
Prob > chi2	0.03	0.11	0.15	0.13	3.32	0.00	0.00	0.00	0.00	9.81	0.00	0.00	0.00	0.00	0.00

P-value in parenthesis. * p=<0.1; ** p<0.05; *** p<0.01