



Ca' Foscari
University
of Venice

**Master's Degree programme
in Global Development and Entrepreneurship**

Final Thesis

**Unexploited export potential in Italian
Prosecco: a gravity analysis of 211 markets,
2018–2025**

Supervisor

Ch. Prof. Lisa Crosato

Graduand

Beatrice Querzoli

Matriculation Number 889028

Academic Year

2025 / 2026

Abstract

This thesis examines Italian Prosecco exports to 211 destinations between 2018 and 2025, with the aim of identifying markets that absorb less than their economic characteristics would predict. The dataset is constructed on the universe of World Bank economies rather than on observed flows and therefore includes 454 zero-flow observations, which existing analyses, restricted to the leading markets, do not consider. The gravity model is estimated by Poisson pseudo-maximum likelihood, which accommodates zero flows and does not require the consistency condition that the heteroskedasticity test rejects for the log-linear specification. The threshold above which a gap between predicted and realised exports becomes interpretable is measured by withholding entire destinations from estimation and observing how far the model misses them. The markets Prosecco has never reached prove to hold no commercial potential: they are worth 0.37 per cent of current exports, and ninety-six per cent of that figure is accounted for by territories whose trade is recorded under another jurisdiction or by markets in which the sale of alcohol is prohibited. Potential is concentrated instead in eleven active but under-penetrated markets, worth 142 million euros a year. A counterfactual in which tariffs are set to zero shows that part of the shortfall is associated with trade barriers rather than with commercial factors: four further markets emerge, and the combined gap rises to 293 million.

Declaration on the use of artificial intelligence tools

The undersigned Beatrice Querzoli, matriculation number 889028, enrolled at Ca' Foscari University of Venice in the Master's Degree programme in Global Development and Entrepreneurship, author of the thesis *Unexploited export potential in Italian Prosecco: a gravity analysis of 211 markets, 2018–2025*, supervisor Prof. Lisa Crosato, aware of the penalties provided for by art. 76 of the Consolidated Act, Presidential Decree no. 445 of 28 December 2000, and of the forfeiture of benefits provided for by art. 75 of the same Consolidated Act in the event of false or untruthful declarations, under her own personal responsibility pursuant to artt. 46 and 47 of Presidential Decree no. 445 of 28 December 2000,

DECLARES

that she has made use of generative artificial intelligence technologies to support certain aspects of the drafting of her thesis.

The tool used is Claude, developed by Anthropic and accessed through the claude.ai interface in the period between April and September 2026. Over that period the tool was updated several times by the provider; in the final stage of the work the version used was Claude Opus 5.

The use of the tool concerned the drafting and revision of certain passages of the text, with the aim of making the exposition clearer and the technical register more appropriate; the organisation of the material into sections; some suggestions of R code, adopted in the course of verifying the results; and support in the collection and checking of bibliographic references.

The construction of the dataset, the extraction and processing of the data, the econometric estimates, the tables and the figures are the exclusive work of the author and were produced in R. No data and no quantitative result reported in the thesis was generated by the artificial intelligence tool.

FURTHER DECLARES

that all material generated by AI has been carefully verified, reviewed and reworked by the author, who assumes full responsibility for the quality, the accuracy and the scientific integrity of the work.

Venice, 30 September 2026

Beatrice Querzoli

Contents

Abstract	i
Declaration on the use of AI tools	ii
Introduction	1
1 Prosecco and the economics of its exports	3
1.1 Prosecco as an object of economic analysis	3
1.2 Prosecco in the world wine market	4
1.3 Previous evidence on Prosecco exports	5
1.4 The gravity framework	7
1.5 Research question and contribution	10
2 The geography of Prosecco exports: dataset construction and descriptive evidence, 2018–2025	12
2.1 Data sources and sample construction	12
2.2 Data cleaning and the treatment of problematic observations	13
2.3 Aggregate dynamics, 2018–2025	16
2.4 The geographical recomposition of export flows	18
2.5 Market concentration	21
2.6 Zero trade flows and the extensive margin	22
2.7 Descriptive statistics of the model variables	24
2.8 The construction of the tariff variable	26
2.9 Summary	28
3 Estimating the gravity model	30
3.1 The empirical model	30
3.2 The choice of estimator	32

3.3	Building the specification	35
3.4	Results	37
3.5	Comparison with the benchmark study	40
3.6	Robustness	41
3.7	Limits of identification	43
3.8	The United States tariff of 2025	44
3.9	Summary	48
4	Unexploited export potential	50
4.1	Measuring potential out of sample	50
4.2	The markets never served	53
4.3	Under-served active markets	55
4.4	The contribution of tariffs	58
4.5	Summary and qualifications	60
	Conclusions	62
	The question and the answer	62
	Implications for the sector	62
	What the analysis cannot establish	63
	Directions for further work	64
	References	66
	Data sources	68
	Software	69

List of Tables

1.1	Variable specification, operational definitions and data sources	8
2.1	Data cleaning interventions and their effect on sample size	13
2.2	Corrections applied to the free trade agreement dummy	15
2.3	Prosecco exports: value, quantity and unit value, 2018–2025	16
2.4	Export value by macro-region: shares and growth, 2018–2025	19
2.5	Top ten destination markets by export value, 2025	20
2.6	Descriptive statistics, 2018–2025 (N = 1688)	25
2.7	Provenance of the tariff variable	28
3.1	Determinants of Prosecco export value under three estimators	32
3.2	Ability of each model to reproduce aggregate exports, 2018–2025	33
3.3	Construction of the specification, PPML on export value	36
3.4	Preferred specification	37
3.5	Replication of Echevarria Garay et al. (2025)	40
3.6	Sensitivity to the exclusion of the leading markets	42
3.7	Within-country variation of the regressors	43
3.8	Italian Prosecco exports to the United States relative to control markets	46
4.1	Out-of-sample predictive performance	51
4.2	The four specifications compared out of sample	52
4.3	The thirty-nine never-served markets by category	54
4.4	Active markets importing less than predicted	55
4.5	Potential under current tariffs and under a zero-tariff counterfactual	58
4.6	The ranking under alternative specifications	59

List of Figures

2.1	Export value and quantity, 2018–2025	17
2.2	Value, quantity and unit value, index numbers (2018 = 100)	18
2.3	Composition of export value by macro-region, 2018–2025	19
2.4	Market shares of the ten leading destinations, 2018 and 2025	21
2.5	Concentration of export flows, 2018–2025	22
2.6	Active markets and zero trade flows by year, 2018–2025	23
2.7	Applied tariff by destination group, 2025	24
3.1	Italian Prosecco exports to the United States, deviation from control markets . .	47
4.1	Predicted against realised Prosecco exports, annual averages 2018–2025, 172 destinations with positive flows.	57

Introduction

Prosecco did not legally exist under that name until 2009. Before the reform of that year the word denoted a grape variety rather than a place, and the appellation as it is now understood — the Denominazione di Origine Controllata covering nine provinces of Veneto and Friuli Venezia Giulia, with the Conegliano Valdobbiadene and Asolo areas at its centre — was created by administrative act. In 2025, seventeen years later, Italy exported some 1.8 billion euros of it. More than four fifths of what the appellation produces leaves the country. That ratio is the starting point of this thesis.

A supply chain selling four bottles abroad for every one sold at home depends on the accessibility of foreign markets in a way that a domestically oriented producer does not, and it depends on conditions over which it exercises no control: the tariffs each destination applies, the distance its product must travel, the willingness of distant consumers to adopt an unfamiliar category and trade policy decided in capitals where the sector has no voice. The United States tariff on European wine imposed in August 2025 illustrated how quickly those conditions can change and how little warning the sector receives. Dependence of that degree makes the geography of sales a strategic question rather than a commercial detail, and the appellation has treated it as one.

The language of diversification is pervasive in the sector's own documents and it is almost always framed in terms of new markets: destinations where Prosecco is not yet present and might be. What has not been established is whether such destinations exist in any commercially meaningful sense.

Answering that requires two things the existing literature does not provide. It requires a model of how much a market of given characteristics would be expected to absorb, so that what a market does import can be compared with something. And it requires trade data covering the destinations where Prosecco does not arrive, which are by construction absent from any dataset built on observed flows. This thesis supplies both.

The dataset assembled in Chapter 2 covers the universe of economies recognised by the World Bank rather than the set of markets Italy has actually supplied: 211 destinations over the years 2018 to 2025, for 1688 country-year observations, of which 454 record no trade at all. The gravity model estimated in Chapter 3 explains exports by the size and income of each destination, its distance from Italy, the alcohol its inhabitants consume, the tariff it applies and the wine it produces domestically. Chapter 4 compares what the model predicts with what each market actually imports and interprets the difference.

Two features of the approach deserve mention at the outset, because they distinguish it from the way such exercises are usually conducted. The first is the choice of estimator. The gravity equation is conventionally estimated by taking logarithms, a transformation that cannot accommodate a zero and would therefore discard a quarter of this sample — precisely the quarter

that the question concerns. Poisson pseudo-maximum likelihood is used instead, and Chapter 3 sets out the grounds for the choice, which rest on the treatment of zeros and on a condition of consistency that the data reject rather than on any comparison of goodness of fit.

The second is the treatment of the residual. Every market imports more or less than a model predicts, and the difference is not by itself evidence of anything: a model that misses a market by half has told you about its own imprecision, not about the market. The criterion by which a gap is retained for examination is therefore measured rather than assumed. Destinations are withheld from estimation in rotation and predicted afterwards, which measures the error the model typically makes on a market it has not seen, a median factor of 2.71, and only gaps exceeding that factor are retained.

The results can be stated in advance, since nothing is gained by withholding them. The markets Prosecco has never reached contain no commercial potential worth the name. All thirty-nine of them together are predicted to absorb 5.16 million euros a year, under four tenths of one per cent of current exports, and the greater part of even that figure is explained by two things that are not opportunities: six territories whose trade is recorded under another customs jurisdiction, where the product already arrives and is counted elsewhere, and thirteen markets in which the sale of alcohol is prohibited or culturally absent. What remains genuinely unexplored is worth 0.21 million euros a year.

The potential lies instead among markets that already import Prosecco and import far less than their characteristics predict. Eleven such destinations are identified, worth a combined 142 million euros a year, a tenth of current exports and twenty-seven times the potential of the unserved set. China, Ireland, South Korea, Japan, Denmark and Brazil head the ranking. A further exercise, in which tariffs are set to zero and everything else held constant, separates the part of each shortfall associated with protection from the part associated with commercial factors, and brings four additional markets into view: Switzerland, which imports almost exactly what the model predicts given a duty of 40.1% and for which the zero-tariff scenario implies three times that level, is the clearest case.

These findings identify markets warranting investigation rather than sales that could be captured. A gravity model observes income, distance, tariffs and drinking habits; it does not observe whether a distributor exists, whether consumers prefer another style of sparkling wine or whether recorded imports are consumed where they are recorded. What it can do is narrow a field of 211 destinations to a dozen that merit examination one by one, which is a contribution of a modest and specific kind.

The thesis is organised as follows. Chapter 1 introduces Prosecco as an object of economic analysis, situates it in the world wine market, reviews the available evidence on its exports and sets out the research question. Chapter 2 documents the construction of the dataset, including two errors in the source extraction that were identified and corrected, and the assembly of the tariff variable, which required verifying the applied rate for each destination individually. Chapter 3 estimates the gravity model, defends the choice of estimator, reports the results and their robustness and examines the effect of the 2025 United States tariff on monthly data. Chapter 4 measures unexploited potential. The conclusions state what the analysis cannot establish.

Chapter 1

Prosecco and the economics of its exports

1.1 Prosecco as an object of economic analysis

In 2025 Italy exported approximately 1.8 billion euros of Prosecco, a wine that did not legally exist under that name until 2009 and that thirty years earlier was little known outside the Veneto. The appellation consistently exceeds 700 million bottles a year and records an export orientation index above 81%: more than four bottles in five leave the country. No other Italian agricultural product has undergone an expansion of comparable speed and few are as exposed to conditions determined beyond the national border.

That exposure is the reason this thesis exists, and it is also what makes Prosecco worth studying on its own rather than as a component of the Italian wine sector. A supply chain selling four bottles abroad for every one sold at home depends on the accessibility of foreign markets in a way a domestically oriented producer does not: on tariffs, on distance, on the willingness of distant consumers to adopt an unfamiliar category and on trade policy decided where the sector has no voice.

To understand the reasons behind this performance it is necessary to deconstruct the Prosecco system, framing it not as a commercial phenomenon but as an integrated socio-economic and institutional model. Within industrial organisation theory the system constitutes a textbook case of a Marshallian industrial district that has evolved into a modern rural competitiveness cluster, in line with the structural paradigm of Porter (1998) and the applied analyses of Barisan et al. (2016). Its competitive strength lies in spatial concentration and in a dense territorial network of heterogeneous but interdependent actors, whose geographical proximity within the Venetian economic fabric generates positive externalities, continuous knowledge spillovers and a substantial reduction in transaction costs.

The efficacy of this cluster was long hindered by a legal vulnerability on the international stage, resolved only through the regulatory reform of 2009. Until that year the term Prosecco identified a grape variety rather than a protected geographical indication and under the intellectual property regime of the TRIPS Agreement the name of a plant variety cannot be exclusively protected. Any producer anywhere could plant the variety and market the result as Prosecco. Meloni and Swinnen (2018) show how this asymmetry exposed Italian production to systematic free-riding and to competition on marginal cost typical of undifferentiated commodities.

The reform altered the legal architecture of the market through a dual toponymic and legal oper-

ation: the grape variety was renamed with its historical synonym Glera, while the term Prosecco was bound exclusively to a geographical indication, divided between the interregional DOC and the DOCG areas of Conegliano Valdobbiadene and Asolo. The product thereby became a club good protected by international intellectual property rights, allowing producers to capture a positional rent derived from territorial exclusivity and mitigating cross-border information asymmetry through a collective reputational standard.

This institutional framework provides the methodological justification for focusing the empirical investigation on Prosecco rather than on Italian wine as a whole. Attempts to map trade flows using national aggregates (Castillo et al., 2016; Dal Bianco et al., 2016) suffer from aggregation bias, since treating Italian wine as a single category obscures profound heterogeneity in demand elasticities, price dynamics and entry barriers across segments. As Pomarici et al. (2021) document, the national segment of traditional still wines faces structural stagnation and declining consumption in mature western markets, while the sparkling segment exhibits counter-cyclical resilience. Prosecco is not a generic component of the sector but the principal engine through which Italian wine exports have expanded internationally.

A further specificity concerns consumer behaviour. Onofri and Boatto (2022), estimating the product life cycle econometrically, document how Prosecco has de-seasonalised the consumption of sparkling wine. Detached from the special-occasion positioning typical of Champagne, it has been recoded as a flexible lifestyle product associated with daily socialising and with the global ritual of the aperitivo. This positioning generates idiosyncratic price dynamics and patterns of substitutability, and requires modelling that isolates the gravity variables correctly rather than assuming the behaviour of wine in general.

Finally, the exclusive focus on Prosecco permits a level of empirical precision that aggregate analysis cannot reach. The investigation uses customs data isolated through the combined nomenclature code CN 2204.10.15, which excludes other sparkling categories and generic wines and yields a clean mapping of transaction costs and market accessibility.

1.2 Prosecco in the world wine market

The market in which Prosecco competes displays a degree of trade integration unique within the global agri-food landscape. According to the International Organisation of Vine and Wine (OIV, 2025), against an estimated world consumption of 214.2 million hectolitres, cross-border exports stand stably around 99.8 million hectolitres, generating a total value of 35.9 billion euros. The ratio between the two yields a trade internationalisation index of 46.6%: nearly one bottle in two uncorked anywhere in the world has crossed at least one customs border before reaching the consumer.

The qualitative composition of that figure reveals a structural asymmetry generated by product differentiation. The global average export price is 3.60 euros per litre, but bottled wine, while accounting for 50.8% of the volume traded, captures 67% of the monetary value, at an average of 4.70 euros per litre. If wine behaved as an undifferentiated commodity, trade would concentrate in the bulk segment in order to minimise logistical costs per unit of value; it does the opposite.

The gap between volume and value introduces a central feature of demand in the sector: premiumisation (del Rey and Mueller Loose, 2025). In mature markets per capita consumption

decouples from the overall value of sales, reflecting an increased willingness to pay among consumers who adopt the philosophy of drinking less but better. Wine is thereby elevated to an experiential and positional good, bound to intangible signals of status, sustainability and territorial identity. In such an environment the country-of-origin effect and the depth of reputational capital act as endogenous quality signals that reduce transaction costs and guide international preferences (Agostino and Trivieri, 2014).

This pattern is what the classical models of agricultural trade fail to accommodate. Under the factor-endowment framework, trade in agricultural goods is governed by comparative advantage derived from relative resource endowments, and a country abundant in the relevant factor emerges as a net exporter while reducing imports of substitutes towards zero. The wine market contradicts that rigidity, offering one of the clearer empirical illustrations of the monopolistic competition with product differentiation formalised by Krugman (1980): trade is driven not by the search for the lowest marginal price of a standardised product but by a structural preference for variety, which generates vigorous two-way flows in which the same countries act simultaneously as importers and exporters of different segments and brands. The point matters here because it justifies treating Prosecco as a differentiated product with a demand of its own, but it is a premise of the analysis rather than its subject.

Market equilibrium rested until the late twentieth century on the dominance of Old World producers, whose competitiveness was safeguarded by viticultural tradition and by rigid institutional frameworks of geographical protection. From the 1990s the emergence of New World producers challenged that structure, introducing varietal standardisation and marketing models that redefined the choice criteria of the global consumer (Mariani et al., 2012). As Anderson and Pinilla (2022) document, the globalisation of the last three decades has not merely intensified trade volumes but reshaped the economic geography and qualitative composition of flows.

This dense network of trade is not free of physical and institutional frictions; it is those frictions that justify the use of gravity models. Dal Bianco et al. (2016) demonstrate that alongside a general reduction in direct customs tariffs there has been a proliferation of non-tariff measures — sanitary protocols, technical barriers, labelling requirements — which operate as asymmetric frictions capable of reshaping export corridors. The desire for variety drives expansion, but the actual geographical distribution of flows remains constrained by transport costs and trade barriers.

1.3 Previous evidence on Prosecco exports

Despite the position Prosecco occupies in the global sparkling wine market, the econometric modelling of its trade flows has received relatively limited attention, the literature preferring consumer surveys (Onofri et al., 2015) or institutional analyses of the district (Pomarici et al., 2019). Studies applying gravity analysis to this appellation specifically remain rare. The closest antecedent to the present work is Echevarria Garay et al. (2025), *Prosecco Under Pressure: Modelling Export Flows and U.S. Tariff Scenarios Using a Gravity Model Approach*. Their research question arises from the need to understand the attracting and friction factors governing Prosecco exports and to evaluate the impact of the customs tariffs announced by the United States administration in July 2025, which threatened to raise duties on European Union agri-food products from 10% to 30%, with escalation to 60% in the event of retaliation.

The authors analyse the years 2017 to 2023 on the forty leading destinations, using the loga-

rithm of export volume in kilograms under CN 2204.10.15 as dependent variable and aggregating the flows of Prosecco DOC, Conegliano Valdobbiadene DOCG and Asolo DOCG. The specification includes the economic mass of exporter and importer measured by GDP, the linear distance between Rome and partner capitals, a continuous measure of applied tariffs in dollars per kilogram, an indicator for the presence of specific duties, a time trend and a pandemic dummy.

Their results attribute a strong attraction effect to economic mass, with an elasticity of 0.644, and a significantly negative effect to distance, at -0.424 , suggesting that logistical costs remain a relevant constraint for Prosecco competitiveness. Ad valorem tariffs are associated with lower flows, at -0.130 , while the indicator for specific duties carries a positive coefficient of 1.394, which the authors interpret as evidence that the predictability of a fixed barrier is less distortive than the volatility of an ad valorem one. The pandemic variable is not statistically significant and the time trend of 0.147 is consistent with structural growth associated with brand consolidation.

Through predictive simulations for 2025 to 2027 the authors estimate that tariffs of 30% and 60% would produce immediate losses of 24.3% and 30.3% respectively against a counterfactual without additional barriers, with volumes recovering to the 2024 level by the end of 2026 in the first scenario and of 2027 in the second. To cushion such shocks they suggest supply-side management by the protection consortia, through the regulation of harvest reserves and the activation of reserve areas.

Two observations should be recorded, since both bear on how that study relates to the present one. The first is that the measure eventually implemented differs from the one simulated: the tariff applied from 1 August 2025 was of 15%, not 30%, so their scenarios remain counterfactuals for a policy that was not adopted. Section 3.8 examines the effect of the measure that was.

The second concerns the sample. The restriction to the forty leading markets is a consequence of the estimator: the log-linear transformation cannot accommodate a zero and confining the analysis to markets that trade in every year avoids the difficulty rather than resolving it. That design is entirely appropriate to their question, which concerns the determinants of established flows; it cannot address a question about markets that lie outside the sample.

Their work sits within a broader literature on frictions in the global wine trade. Dal Bianco et al. (2016) show that tariff and non-tariff barriers significantly reduce wine flows, with the effect partially mitigated for protected designation products. Agostino and Trivieri (2014) find that the reputation associated with geographical indications acts as a trade facilitator, diminishing the negative impact of cultural and geographical distance. On exogenous shocks, Macedo et al. (2023) confirm that premium wine exports from major European producers displayed unusual structural stability during the pandemic, a behaviour also observed in Prosecco specifically.

Their specification also includes the gross domestic product of Italy, which returns a coefficient of -0.084 with a p-value of 0.959. The result is structural rather than accidental: with a single exporter, the exporter's own income varies only over time and is therefore nearly collinear with the time trend included in the same equation. Section 3.1 addresses the same difficulty by absorbing all exporter-side variation into year fixed effects.

1.4 The gravity framework

Translating these dynamics into an estimable form requires a methodology capable of relating export flows to transaction costs, tariff barriers and macroeconomic factors. The instrument adopted is the gravity model, introduced by Tinbergen (1962) and subsequently equipped with microeconomic foundations, most influentially by Anderson and Van Wincoop (2003), who derived the equation from a general equilibrium model of trade in differentiated goods and showed that bilateral flows depend not only on the frictions between two countries but on the trade costs each faces with all its other partners. That result has a direct consequence for the design adopted here, set out in Section 3.1: with a single exporter the multilateral resistance terms cannot be included alongside distance, contiguity and the landlocked indicator, since destination fixed effects would absorb them.

In its canonical formulation, borrowed by analogy from Newton's law of universal gravitation, the model postulates that trade between two partners is proportional to the product of their economic masses and inversely proportional to the distance separating them, which serves as a proxy for transport and transaction costs. In the present case the structure of the data dictates an important specification. Because the investigation concerns exports of a single Italian product, the model is of the single-exporter type: the origin is constantly Italy and the variability of the dataset is driven entirely by the characteristics of the destinations and by the temporal dimension. The economic mass of the exporting country and its structural characteristics do not vary across destinations and are absorbed by the intercept and by year fixed effects.

The multiplicative specification estimated in Chapter 3 takes the following form:

$$X(j,t) = \exp[\beta_0 + \beta_1 \ln GDP(j,t) + \beta_2 \ln POP(j,t) + \beta_3 \ln Dist(j) + \Gamma'Z(j) + \Phi'W(j,t) + \delta(t)] \cdot \varepsilon(j,t)$$

where $X(j,t)$ is the flow of Prosecco to destination j in year t , GDP and POP the gross domestic product and population of the destination, and Dist the bilateral distance from Italy. The vector $Z(j)$ collects the time-invariant destination characteristics — contiguity with Italy, absence of sea access, per capita alcohol consumption and domestic wine production per head — and $W(j,t)$ the time-varying institutional and policy variables, namely European Union membership, the existence of a preferential trade agreement and the applied ad valorem tariff. The term $\delta(t)$ is a full set of year fixed effects and $\varepsilon(j,t)$ a multiplicative error satisfying $E[\varepsilon | \cdot] = 1$, which is the only distributional assumption the estimator requires.

Bilateral distance is taken from the CEPII GeoDist database (Mayer and Zignago, 2011) for the 197 destinations it covers; for the remaining fourteen, mostly micro-territories absent from that source, the great-circle distance from Rome was computed directly, as documented in Chapter 2. Contiguity and the landlocked indicator come from the same database, the first capturing the reduced road transport costs of a shared border and the second the higher costs of intermodal transport imposed on destinations without direct port access.

Three variables considered at the design stage do not appear in the specification and the reasons differ. A common-language indicator was dropped because Italian is an official language in no destination outside Italy and the available variants identify too few cases to sustain estimation. Indicators for the pandemic and post-pandemic years were dropped because they are exact functions of the calendar year and therefore perfectly collinear with the year fixed effects, which perform the same role less restrictively: rather than imposing a common shift on 2020 and 2021, they allow each year its own level and absorb the pandemic together with exchange

rate movements, freight costs and everything else common to all destinations within a year. The pandemic indicator reappears only in Section 3.5, where the benchmark specification being replicated uses a time trend and a Covid dummy in place of year fixed effects.

Three variables absent from the canonical gravity equation are included instead, for reasons specific to the product. Per capita alcohol consumption captures a propensity that varies across destinations by more than an order of magnitude for cultural and legal reasons which income and distance do not register, and Section 3.3 shows that without it the model fails a functional-form test. The applied tariff enters because the product faces duties ranging from zero to 150% across the sample. Domestic wine production per head enters because a destination with its own wine industry consumes a close substitute, and Section 4.4 shows that omitting it places Spain, Portugal and Chile at the head of the ranking of under-served markets.

On the estimation front, the log-linear transformation of the specification has represented standard methodological practice for decades. As demonstrated by Silva and Tenreyro (2006), however, ordinary least squares applied to the logarithmic form suffers from two difficulties in this setting. The first is that zero trade flows — a frequent occurrence when analysing disaggregated customs data for a single product exported to small markets — make the logarithmic transformation mathematically impossible, which forces their exclusion from the sample. The second is that in the presence of heteroskedasticity, where the variance of the error correlates with the levels of the explanatory variables, log-linearisation yields inconsistent estimates, since consistency then requires a condition on the expectation of the log error that heteroskedasticity generally violates.

This study therefore estimates the equation by Poisson pseudo-maximum likelihood, retaining the log-linear specification only as a point of comparison. The estimator accommodates zero flows directly, since it works on the multiplicative form without transformation and remains consistent under any form of the conditional variance provided the conditional mean is correctly specified. Chapter 3 sets out the grounds for the choice in detail and Chapter 4 compares the predictive accuracy of the two estimators on destinations withheld from estimation.

Table 1.1: Variable specification, operational definitions and data sources

Type	Variable	Symbol (dataset name)	Definition	Sign	Source
Dependent	Prosecco exports	$X(j,t)$ — export_ value_euro, export_mass_ kg	Annual flow from Italy to destination j in year t , in euros and in kilograms. Customs code CN 2204.10.15	—	ISTAT Coeweb
Economic mass	Partner GDP	$GDP(j,t)$ — gdp_partner	Gross domestic product of destination j , current USD	+	World Bank WDI
Demographic mass	Population	$POP(j,t)$ — pop_partner	Total population of destination j ; jointly with GDP it identifies the elasticity to income per head	+/-	World Bank WDI

Table 1.1: Variable specification, operational definitions and data sources (*continued*)

Type	Variable	Symbol (dataset name)	Definition	Sign	Source
Trade cost	Bilateral distance	Dist(j) — dist	Distance in kilometres between Italy and destination j	—	CEPII GeoDist; own calculation for 14 territories
Trade cost	Contiguity	Contig(j) — contig	Equal to 1 if j shares a land border with Italy	+	CEPII GeoDist
Trade cost	Landlocked	Landlocked(j) — landlocked	Equal to 1 if j has no sea access	—	CEPII GeoDist
Demand	Alcohol consumption	Alc(j) — alcohol_pc	Litres of pure alcohol consumed per inhabitant per year	+	WHO Global Health Ob- servatory
Supply substitution	Domestic wine production	Wine(j) — wine_pc10	Domestic wine output in tens of litres per inhabitant	—	OIV (2026) and OIV database
Trade policy	Applied tariff	$\ln(1+\tau(j,t))$ — $\ln_tau$	Logarithm of one plus the ad valorem equivalent duty applied to CN 2204.10.15	—	ITC Market Access Map; WITS- TRAINS; DG Trade
Institutional	EU membership	EU(j,t) — EU	Equal to 1 if j is a member of the European Union in year t	+/—	European Commis- sion
Institutional	Free trade agreement	FTA(j,t) — FTA	Equal to 1 if a preferential agreement between the EU and j is in force in year t	+	DG Trade
Time effects	Year fixed effects	$\delta(t)$	Full set of year indicators, absorbing inflation, exchange rates, the pandemic and any other shock common to all destinations	—	Own elaboration

The expected sign of the European Union indicator is left uncertain: Section 3.4 shows that, once tariffs and trade agreements are controlled for, what remains attached to it is a residual without natural interpretation. The names in the third column are those of the variables in the analysis dataset. Those reported as transformations,

ln_tau and wine_pc10, are constructed in the data preparation script from the raw columns tariff_ave and the OIV production series, and do not appear under those names in the dataset file itself. Source: own elaboration.

1.5 Research question and contribution

The purpose of this thesis is to establish where Italian Prosecco exports fall short of what the characteristics of foreign markets would lead one to expect and to quantify the shortfall. The question is prospective rather than descriptive: it concerns not the destinations that already absorb the product but those that, on the evidence of their income, their proximity, their drinking habits and the barriers they impose, might absorb more of it than they do.

Answering it requires three things, and the structure of the thesis follows from them. It requires trade data covering the whole set of possible destinations rather than the markets already served, which Chapter 2 assembles for 211 partners over the years 2018 to 2025. It requires a model of the expected level of trade, which Chapter 3 estimates. And it requires a way of distinguishing a genuine shortfall from the ordinary imprecision of any such model, which Chapter 4 provides by withholding destinations from estimation and measuring how far the model misses them.

The relationship with the study described in Section 1.3 is one of complementarity rather than competition. Echevarria Garay and colleagues ask which determinants explain the magnitude of flows towards markets where Prosecco is already established and how those flows would respond to a tariff; their design is appropriate to that question. The question asked here cannot be posed on their sample, because it concerns markets that lie outside it. Section 3.5 replicates their specification on the data assembled for this thesis and obtains estimates that coincide with theirs closely, which is offered as evidence that the two datasets are consistent rather than as a comparison of merit.

The contribution of this work can be stated in three parts. The first is the extension of the sample to the universe of destinations. Building the panel on the complete set of World Bank economies rather than on observed flows admits 39 markets that have never imported Prosecco and 454 country-year observations of zero trade, which amount to a quarter of the panel. This requires an estimator capable of accommodating them. The extension is not an improvement in technique for its own sake: without it the question posed above cannot be asked at all, since the markets it concerns would be absent from the data.

The second is the calibration of the notion of potential. The comparison between predicted and realised exports produces a residual for every market and a residual is not by itself evidence of an opportunity. Chapter 4 measures, by predicting destinations the model has not seen, the size of the error the model typically makes on an unfamiliar market, and uses that error as the criterion for retaining a gap rather than a cut-off chosen for convenience. The exercise also delivers a negative result, since it indicates that the markets Prosecco has never reached are worth under half of one per cent of current exports and that the potential lies almost entirely among destinations that already import the product in quantities below what their characteristics suggest.

The third concerns trade policy. Tariffs enter the model as an explanatory variable, which permits a counterfactual in which they are removed while all other characteristics are held constant, and thereby a separation between the part of each shortfall associated with protection and the part associated with commercial factors. The distinction matters because the two call for different responses. Chapter 3 also examines the effect of the United States tariff of 2025

directly, on monthly data, which provides evidence on a policy change rather than an inference from a cross-sectional elasticity.

Two limitations should be stated at the outset, since they bear on how the results may be read. The model explains trade by the characteristics it observes and a market may import less than predicted because it prefers a different product, because no distributor has established a position there, or simply because the prediction is wrong; the analysis identifies markets warranting investigation and does not establish the cause of the anomaly. And with a single exporter, as noted in Section 1.4, the multilateral resistance terms cannot be included alongside the trade-cost variables of interest, so the coefficients on distance and contiguity are not directly comparable with estimates from multi-country specifications.

Chapter 2

The geography of Prosecco exports: dataset construction and descriptive evidence, 2018–2025

2.1 Data sources and sample construction

The empirical analysis rests on an original panel dataset assembled for this research, covering Italian exports of Prosecco to 211 partner countries and territories over the period 2018–2025. This section documents the construction of the dataset, the treatment applied to problematic observations and the resulting sample, so that every figure reported in the remainder of the thesis can be traced back to an identified source and a documented processing step.

Trade flows are drawn from the Coeweb database of the Italian National Institute of Statistics (ISTAT), which publishes foreign trade statistics disaggregated at the eight-digit level of the Combined Nomenclature. Prosecco is isolated through code CN 2204.10.15, which covers sparkling wine with a protected designation of origin bearing the name Prosecco and therefore aggregates the three denominations of the appellation system: Prosecco DOC, Conegliano Valdobbiadene Prosecco DOCG and Asolo Prosecco DOCG. The annual series was extracted in August 2026 and covers the calendar years 2018 through 2025.

Two dependent variables are retained. The first is the export value, expressed in current euros, which is the natural unit for a trade-policy analysis concerned with the revenue exposure of the supply chain. The second is the quantity exported, which ISTAT reports as net mass in kilograms. Their ratio yields the unit value in euros per kilogram, used throughout as a proxy for the positioning of the product in each destination market.

The unit of quantity deserves an explicit note, because it is a common source of error in this literature. ISTAT publishes both the net mass in kilograms and a supplementary unit in litres, and the two are numerically close for wine, whose density is near one kilogram per litre, but they are not the same variable. The series used here is the net mass. Where a per-litre conversion is required — as it is for the United States, whose most-favoured-nation duty on sparkling wine is levied per litre — the conversion is performed explicitly and the assumed density is declared. Reporting value and quantity together is not redundant: as Section 2.3 shows, the two series diverge in a systematic and economically meaningful way in precisely those years in which the sector was hit by an external shock, and a specification that observes only one of them cannot

separate the two channels.

The choice of ISTAT as the primary source, rather than the mirrored import statistics of TradeMap used by Echevarria Garay et al. (2025), reflects a preference for the declarations of the exporting country. Italian customs records are the direct administrative trace of the shipment and avoid the reconciliation problems that arise when partner-country import declarations are used, notably the attribution of goods that transit through a logistics hub before reaching final consumption.

Explanatory variables are drawn from established international sources. Partner gross domestic product at current prices, expressed in current United States dollars, and total population come from the World Bank World Development Indicators; figures for 2025 are provisional at the time of extraction. Since the specification is logarithmic and the exporter is constant, the currency in which partner GDP is denominated affects only the intercept and not the estimated elasticities. Bilateral distance, contiguity and the landlocked indicator come from the CEPII GeoDist database (Mayer and Zignago, 2011), with distance measured as the great-circle distance between Rome and the principal agglomeration of the partner country. European Union membership is coded from the accession and withdrawal dates published by the European Commission. The free trade agreement dummy is coded from the register of agreements in force maintained by DG Trade, cross-checked against EUR-Lex and the Access2Markets portal for the entry-into-force date of each individual agreement.

Per capita alcohol consumption is taken from the World Health Organization Global Health Observatory. Domestic wine production is drawn from the International Organisation of Vine and Wine: for the twenty producers above one million hectolitres from the country table of the State of the World Wine Sector in 2025 (OIV, 2026), and for nine smaller producers relevant to the analysis from the OIV statistical database, extracted in August 2026. Section 3.1 documents the construction of the production variable, which enters the estimation rather than the descriptive analysis and is time-invariant by construction.

2.2 Data cleaning and the treatment of problematic observations

The raw extraction returned 1369 country-year records. Six categories of problem were identified and addressed before any statistical work was undertaken. Each intervention is recorded in an audit log stored alongside the dataset, reporting the rule applied, the observations affected and the source used for verification.

Table 2.1: Data cleaning interventions and their effect on sample size

Step	Issue identified	Treatment applied	Affected	Remaining
0	Raw extraction from ISTAT Coeweb	/	/	1369

Table 2.1: Data cleaning interventions and their effect on sample size (*continued*)

Step	Issue identified	Treatment applied	Affected	Remaining
1	The United States appeared four times per year: the ISO mapping had collapsed four levels of the Coeweb territorial hierarchy onto a single country code	Country-level record selected and validated against two arithmetic identities	24	1345
2	Quantity variable labelled as litres but containing net mass in kilograms	Variable renamed; per-litre conversions performed explicitly where required	/	1345
3	Bilateral distance recorded as zero for six countries, owing to a failed match on the ISO code during the merge	Recomputed as great-circle distance from Rome to the principal agglomeration	48	1345
4	Free trade agreement dummy misreported for four partners	Recoded against the DG Trade register (see Table 2.2)	20	1345
5	Common-language dummy took value one for a single country	Variable dropped: the coefficient would have been identified by one cross-sectional unit	/	1345
6	Nine micro-territories and two economies not covered by World Bank statistics lacked GDP or population	Removed; the excluded flows amount to 0.014% of total export value over the period	68	1688

Source: own elaboration on ISTAT Coeweb, World Bank WDI, CEPII GeoDist, DG Trade, ITC Market Access Map and WHO Global Health Observatory data.

The first intervention is the one with material consequences for every figure in this thesis, and it warrants a full account. In the raw extraction the United States was represented by four separate records in each year. These were not product breakdowns but four nested levels of the

Coeweb territorial classification, which the ISO country mapping had assigned indiscriminately to the same code. The four levels are, in ascending order of aggregation: Central and South America; the United States proper; North America, equal to the United States plus Canada; and the American continent as a whole.

The record initially retained was the largest, which is the continental aggregate. Retaining it would have overstated the single most important market by between 13.5% and 15.9% in every year of the panel, and would simultaneously have counted Canada, Latin America and the Caribbean twice — once inside the American aggregate and once as individual partners — for a cumulative double count of 435.3 million euros over the period. The correct record was identified through two arithmetic identities, namely that the continental total equals North America plus Central and South America, and that North America equals the United States plus the separately recorded Canadian flow. Both identities hold exactly in all eight years. A cross-check of this kind is inexpensive, and is recommended as standard practice whenever a hierarchical trade database is queried, since an aggregation error of this type is invisible in the totals and distorts precisely the largest market.

The recoding of the free trade agreement dummy is reported in Table 2.2, since the variable is one of the institutional determinants on which this thesis places particular emphasis and any error in its coding would translate directly into a biased estimate of the coefficient on European trade agreements.

Table 2.2: Corrections applied to the free trade agreement dummy

Partner	Recoded	Agreement and entry into force
New Zealand	1 from 2024	EU–New Zealand Free Trade Agreement, in force 1 May 2024; removes duties on EU agri-food exports including wine
Kenya	1 from 2024	EU–Kenya Economic Partnership Agreement, in force 1 July 2024; liberalisation of the Kenyan market is asymmetric and phased
Egypt	1 throughout	EU–Egypt Euro-Mediterranean Association Agreement, in force since 2004 and therefore covering the whole sample period
Seychelles	1 throughout	EU–ESA interim Economic Partnership Agreement, applied since 2012 and therefore covering the whole sample period

Source: DG Trade register of agreements in force, cross-checked against EUR-Lex and the European Commission Access2Markets portal, consulted August 2026.

A further issue concerns the balance of the panel. Twelve partners are absent from the extraction in some years, for a total of 27 missing country-year cells. These are without exception marginal destinations that record zero or negligible flows in the adjacent years, and the absence reflects the fact that Coeweb returns a record only when a country appears in the territorial register for that year. The economically correct reading of these cells is a zero flow rather than a missing observation. This is an assumption and is declared as one; the sensitivity of the results to the alternative treatment is verified in Chapter 3.

The final panel comprises 1688 observations on 211 partner countries and territories over eight

years, of which 454 records, equal to 26.9% of the sample, are zero trade flows. These zeros are retained rather than discarded. This is the single most consequential methodological choice of the dataset: a log-linear specification would have required their elimination, reducing the analysis to the set of markets in which Prosecco has already established itself and making the estimation of untapped export potential logically impossible.

Two further errors were identified by cross-checking the partner list against UN Comtrade mirror statistics for HS 220410. The series recorded under the Russian Federation in earlier versions of this dataset in fact held Belarusian values, matching the Coeweb series for Belarus exactly in every year from 2018 to 2025; Russia itself, worth 74.1 million euros in 2025 and the fifth-largest destination market, was absent from the dataset entirely. Thailand was likewise absent throughout. Both omissions trace to country names as rendered in the Italian-language source, Federazione russa rather than Russia and Tailandia rather than Thailandia, which an earlier name-matching step did not recognise. The extraction was repeated from source and every remaining country verified by matching its eight-year value signature against the previous vintage: 162 of 163 matched exactly.

Aggregate values were validated against the same mirror statistics. For 2024, Prosecco accounts for 72.6 per cent of the value and 70.7 per cent of the weight of total Italian sparkling wine exports under HS 220410, proportions consistent with the appellation's share of national sparkling wine production and stable across the period. Ten territories were excluded from the panel for want of national accounts data, their combined trade amounting to 1.88 million euros over the period, or 0.017 per cent of the total; a gravity model cannot use an observation without economic mass.

A second extraction was carried out at monthly frequency, covering the twenty-one leading destinations from January 2020 to May 2026 for 1617 country-month observations, and is used in Section 3.8 to identify the effect of the 2025 United States tariff. Summing the monthly figures reproduces the annual series exactly, to the euro, for every one of the 126 country-year cells over which the two extractions overlap. Since the two were retrieved independently, the agreement provides an external check on the annual panel.

2.3 Aggregate dynamics, 2018–2025

Over the eight years covered by the panel, the value of Italian Prosecco exports rose from 915.7 million euros to 1772.5 million euros, an increase of 93.6% corresponding to a compound annual growth rate of 9.9%. Quantities grew from 222.3 to 420.0 thousand tonnes, an increase of 88.9% and a compound annual rate of 9.5%. The near-coincidence of the two growth rates is itself informative: over the period as a whole, expansion was almost entirely a quantity phenomenon, with the average unit value rising only from 4.12 to 4.22 euros per kilogram, a compound annual rate of 0.3%. Measured against euro-area consumer price inflation over the same period, this represents a substantial erosion of the real unit value.

Table 2.3: Prosecco exports: value, quantity and unit value, 2018–2025

Year	Value	Δ %	Quantity	Δ %	Unit value	Δ %
	(EUR mn)		(kt)		(EUR/kg)	
2018	915.7	/	222.3	/	4.12	/

Table 2.3: Prosecco exports: value, quantity and unit value, 2018–2025 (*continued*)

Year	Value (EUR mn)	Δ %	Quantity (kt)	Δ %	Unit value (EUR/kg)	Δ %
2019	1052.8	+15.0	267.2	+20.2	3.94	−4.4
2020	1009.2	−4.1	273.3	+2.2	3.69	−6.2
2021	1311.4	+29.9	353.3	+29.3	3.71	+0.5
2022	1577.0	+20.3	370.9	+5.0	4.25	+14.6
2023	1640.7	+4.0	359.4	−3.1	4.57	+7.4
2024	1805.1	+10.0	409.1	+13.8	4.41	−3.4
2025	1772.5	−1.8	420.0	+2.6	4.22	−4.3
CAGR	+9.9%		+9.5%		+2.5%	

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15, extracted August 2026. Quantity is net mass.

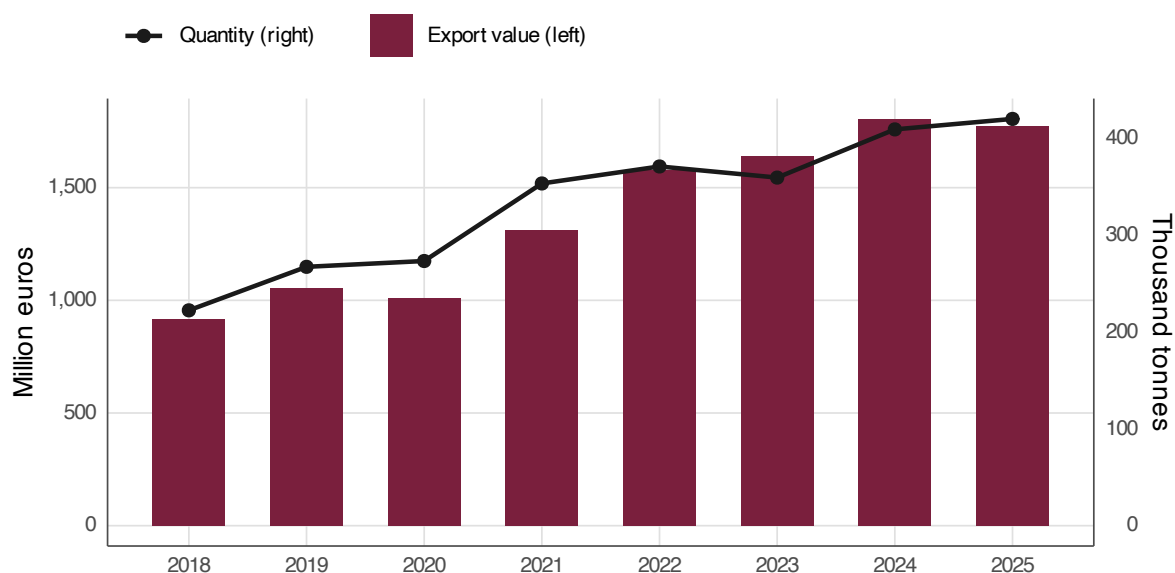


Figure 2.1: Export value and quantity, 2018–2025

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15.

The aggregate trajectory is not a smooth trend, and its two interruptions are the most analytically interesting feature of the series. The first occurs in 2020, when export value falls by 4.1% while quantity continues to grow by 2.2%. A pandemic that had reduced demand would have depressed both series; a contraction confined to value, accompanied by expanding quantities, indicates instead a recomposition of distribution channels. The closure of the hotel, restaurant and catering channel across the principal destination markets shifted sales towards retail and e-commerce, where the same physical product is traded at a materially lower unit value. The unit value falls by 6.2% in that year, the sharpest annual decline in the series.

The disaggregation of the 2020 shock by destination confirms this reading and refines it. The contraction is concentrated almost entirely in the two markets with the highest exposure to on-trade consumption: the United Kingdom, where value falls by 11.8% against a quantity decline of only 4.3%, and the United States, where value falls by 8.6% against a quantity decline of 3.1%. In continental Europe the direction is reversed: German export value rises by 6.8% and

quantity by 17.5%, Belgian value by 11.4% and quantity by 16.0%, French value by 3.6% and quantity by 16.8%. The pandemic did not contract the demand for Prosecco; it redistributed it across channels and across markets, and the aggregate figure conceals both movements.

This finding has a direct bearing on the specification adopted in Chapter 3. Echevarria Garay et al. (2025) estimate a COVID-19 coefficient that is statistically insignificant, with a point estimate of 0.051, and interpret it as evidence of the resilience of Prosecco demand. The evidence assembled here supports that conclusion while identifying its mechanism: a specification in quantity, as theirs is, records no shock precisely because the shock operated on the unit value and not on the quantity traded. Estimating the model on both dependent variables, as this thesis does, allows the two channels to be separated rather than conflated.

The second interruption occurs in 2025, and reproduces the same asymmetry in a different setting: export value falls by 1.8% while quantity grows by 2.6%, and the unit value drops by 4.3%. In the United States, the market directly affected by the tariff measures introduced in the second half of that year, value falls by 3.5% while quantity is essentially unchanged at plus 0.8%. The pattern is consistent with the tariff being absorbed within the distribution margin rather than passed through to quantities over so short a horizon. The observation window closes only five months after the measure took effect, and the point is treated here as descriptive. ISTAT treats the most recent year as provisional and revises it as late returns are incorporated, so the magnitude of the 2025 contraction may change; the figures reported here were extracted in August 2026.

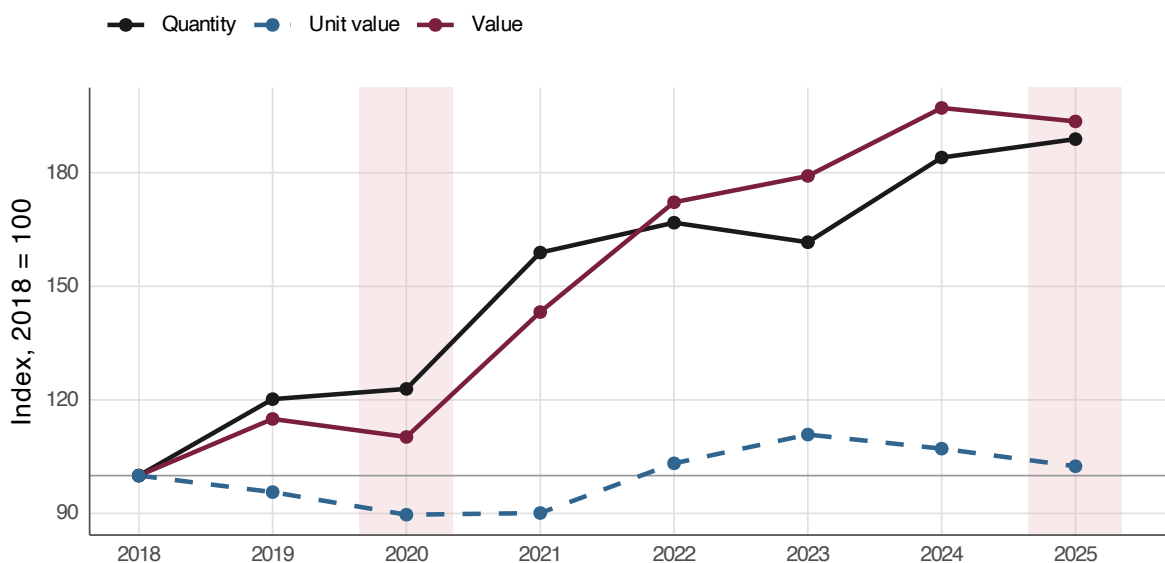


Figure 2.2: Value, quantity and unit value, index numbers (2018 = 100)

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15.

2.4 The geographical recomposition of export flows

Aggregate growth conceals a redistribution of flows across destinations that is more pronounced than the headline figures suggest and that constitutes, for the purposes of this thesis, the central descriptive finding of the period.

Table 2.4: Export value by macro-region: shares and growth, 2018–2025

Macro-region	2018 (mn €)	2025 (mn €)	2018 share	2025 share	Δ share (p.p.)	CAGR
Non-EU	406.7	557.4	44.4%	31.4%	−13.0	+4.6%
Europe						
EU-27	229.3	594.1	25.0%	33.5%	+8.5	+14.6%
North America	231.4	531.2	25.3%	30.0%	+4.7	+12.6%
Asia-Pacific	36.5	59.4	4.0%	3.4%	−0.6	+7.2%
Middle East & North Africa	3.8	12.9	0.4%	0.8%	+0.3	+19.2%
Latin America & Caribbean	5.5	12.2	0.6%	0.7%	+0.1	+12.0%
Sub-Saharan Africa	2.3	5.1	0.3%	0.3%	+0.0	+11.8%
Total	915.7	1772.5	100.0%	100.0%	/	+9.9%

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15. Regional aggregation by the author; the United Kingdom is classified with non-EU Europe throughout, Mexico with North America.

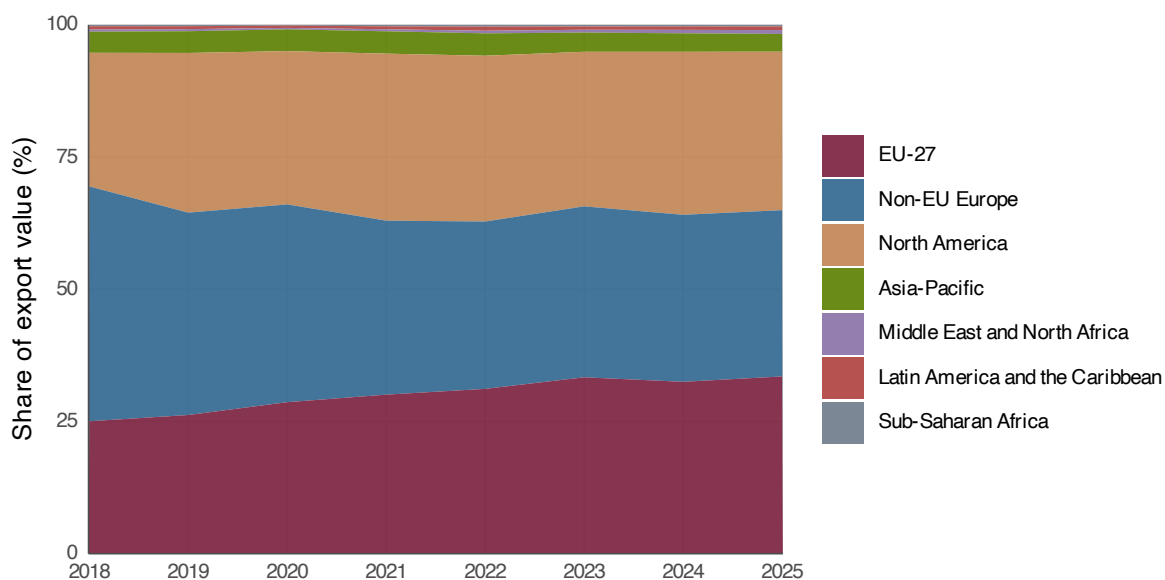


Figure 2.3: Composition of export value by macro-region, 2018–2025

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15.

Non-EU Europe loses 13.0 percentage points of share over the period, falling from 44.4% to 31.4% of export value. The movement is almost entirely accounted for by the United Kingdom,

which alone represented 36.7% of Italian Prosecco exports in 2018 and represents 21.0% in 2025. The British market did not contract in absolute terms — its export value rises from 336.3 to 372.3 million euros — but it grew at a compound annual rate of 1.5% against 9.9% for the aggregate, and this differential is sufficient to displace it from first to second position among destination markets. The interpretation of this stagnation, and specifically the extent to which it is associated with the change in the trade regime following the United Kingdom’s withdrawal from the European Union, is an empirical question that the gravity specification of Chapter 3 is designed to address, since the country transitions from EU membership to a preferential agreement within the sample period.

The share released by the British market is absorbed by the EU-27, whose weight rises from 25.0% to 33.5% at a compound annual rate of 14.6%, and by North America, which rises from 25.3% to 30.0%. Within the European Union the growth is broad-based but strongly differentiated: France records a compound annual rate of 20.7% and Poland of 18.6%, against 9.4% for Germany. The French figure is large enough to warrant caution in interpretation, since part of the flow may reflect logistical re-export through French distribution platforms rather than final consumption. The distinction cannot be resolved with export data alone and is taken up in Chapter 4, where mirrored import data from UN Comtrade are used to bound the absorption capacity of each market.

Outside the traditional perimeter, the fastest-growing regions are also the smallest. The Middle East and North Africa grows at a compound annual rate of 19.2%, Latin America and the Caribbean at 11.7%, sub-Saharan Africa at 11.8%; taken together the three regions account for 1.7% of export value in 2025, against 1.3% in 2018. At the level of individual markets, growth from a low base is considerable: Belarus records a compound annual rate of 51.7%, Georgia of 43.1%, Kenya of 40.4%, Ukraine of 38.4% and Turkey of 36.1%. These rates are computed on small initial values and should not be extrapolated, but they identify the geographical direction in which the extensive margin of the appellation is currently moving, and they anticipate the set of markets on which the export potential analysis of Chapter 4 concentrates.

Table 2.5: Top ten destination markets by export value, 2025

Market	Value (mn €)	Quantity (kt)	Share 2025	Share 2018	Unit value	CAGR
United States	473.7	107.1	26.7%	22.4%	4.42	+12.7%
United Kingdom	372.3	96.5	21.0%	36.7%	3.86	+1.5%
France	135.3	36.6	7.6%	4.0%	3.70	+20.7%
Germany	90.3	20.1	5.1%	5.3%	4.49	+9.4%
Russian Federation	74.1	19.4	4.2%	1.7%	3.82	+25.2%
Belgium	72.8	19.4	4.1%	3.1%	3.75	+14.4%
Switzerland	57.0	11.5	3.2%	4.2%	4.95	+5.8%
Canada	50.2	9.0	2.8%	2.7%	5.58	+10.6%
Austria	46.4	9.7	2.6%	2.0%	4.76	+14.2%
Poland	42.2	10.9	2.4%	1.4%	3.89	+18.6%

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15. Unit value in euros per kilogram; CAGR computed over 2018–2025.

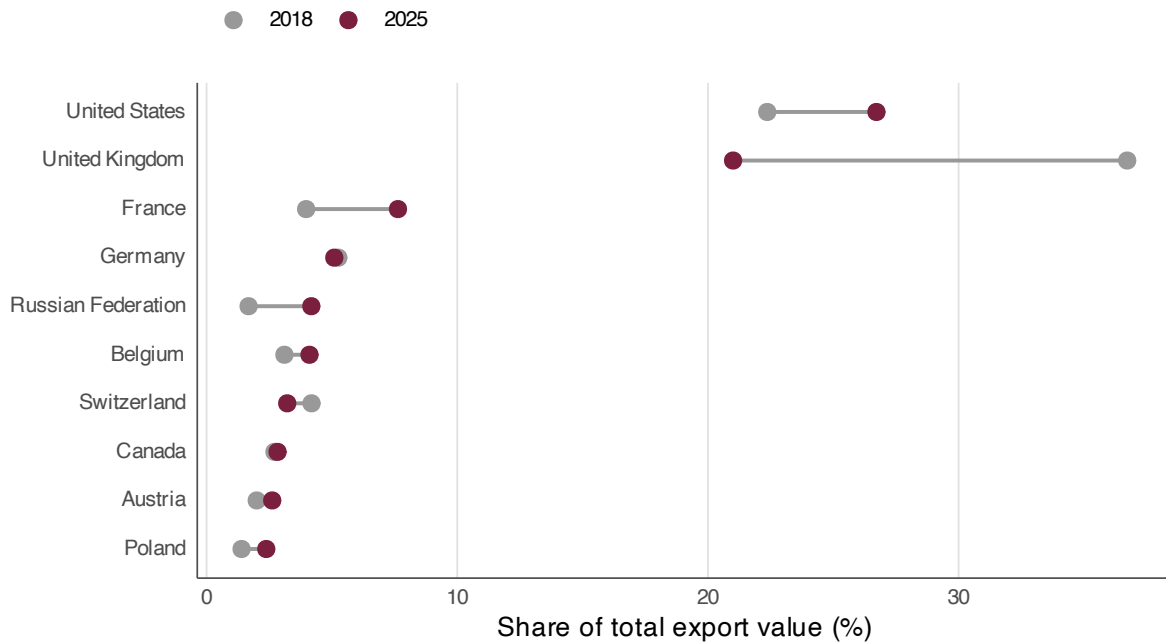


Figure 2.4: Market shares of the ten leading destinations, 2018 and 2025

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15.

The dispersion of unit values across the leading markets — from 3.74 euros per kilogram in France to 5.57 in Canada, a range of nearly fifty per cent — is a first indication that the same physical product occupies materially different positions in different destinations. Since the gravity specification treats the trade flow as a homogeneous quantity, this heterogeneity is absorbed into the error term in a quantity specification and into the dependent variable in a value specification, which is a further reason for estimating the model on both.

2.5 Market concentration

The redistribution documented in the preceding section translates into a measurable, if gradual, reduction in the concentration of export flows. The Herfindahl-Hirschman index computed on destination shares of export value falls from 1953 in 2018 to 1320 in 2025, and the share of the five leading markets from 72.5% to 64.6%. Over the same period the number of markets recording a positive flow rises from 149 to 156, with a maximum of 160 in 2024.

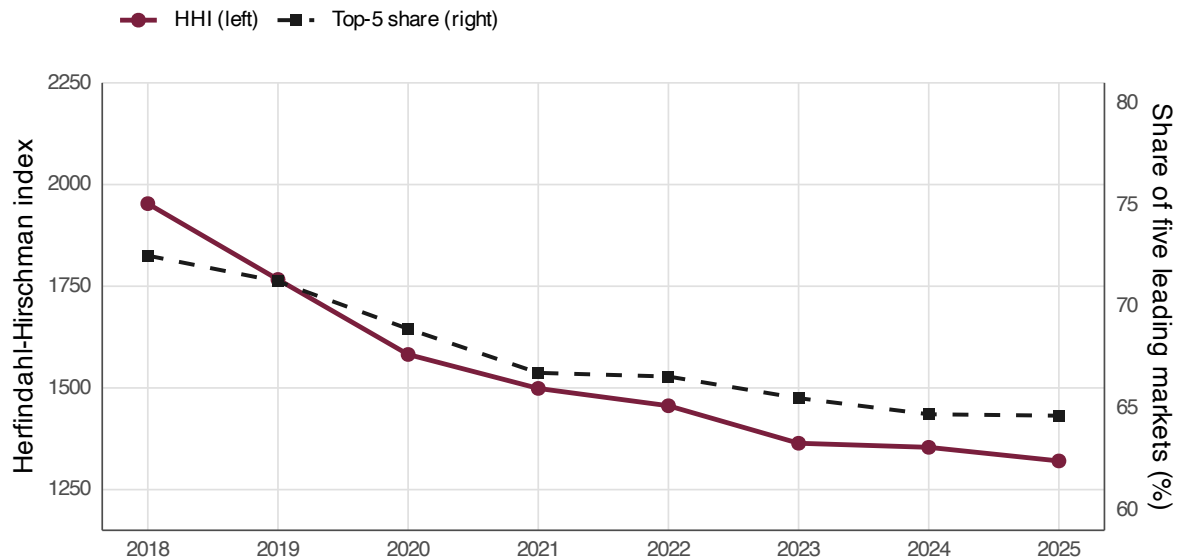


Figure 2.5: Concentration of export flows, 2018–2025

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15. HHI computed on destination shares of export value, scale 0–10000.

The direction of travel is unambiguous but the level remains high. On the thresholds conventionally applied in industrial organisation, a value of 1320 still denotes moderate concentration, and in 2025 the United States and the United Kingdom together absorb 47.7% of export value. Diversification, in other words, has thus far taken the form of a reallocation among mature markets — chiefly from the United Kingdom towards the European Union — rather than a genuine widening of the geographical base. The compound annual growth rates reported in Table 2.4 make the point precisely: the regions growing fastest in relative terms contribute almost nothing in absolute terms, because the base from which they grow is negligible.

It is worth observing that most of the decline in concentration occurs in the first three years of the panel, with the index falling from 1953 to 1499 between 2018 and 2021 and then drifting slowly to 1320. The early movement reflects the relative stagnation of the British market rather than the emergence of new destinations, and the subsequent flattening indicates that this particular source of diversification has largely exhausted itself. This is the empirical premise of the analysis developed in the second part of the thesis. If the reduction in concentration observed to date is essentially a substitution among established destinations, then the exposure of the supply chain to a shock affecting one of its principal markets has not materially diminished, and identifying markets in which the appellation is under-represented relative to what its economic and geographical fundamentals would predict acquires an operational significance that goes beyond the descriptive exercise.

2.6 Zero trade flows and the extensive margin

The panel contains 454 zero trade flows, equal to 26.9% of observations, distributed across 77 partner countries. Thirty-nine of these recorded no exports in any year of the period, and a further thirty-eight were served intermittently. The distinction matters: the first group represents an absolute margin the appellation has never crossed, the second a margin it crosses irregularly.

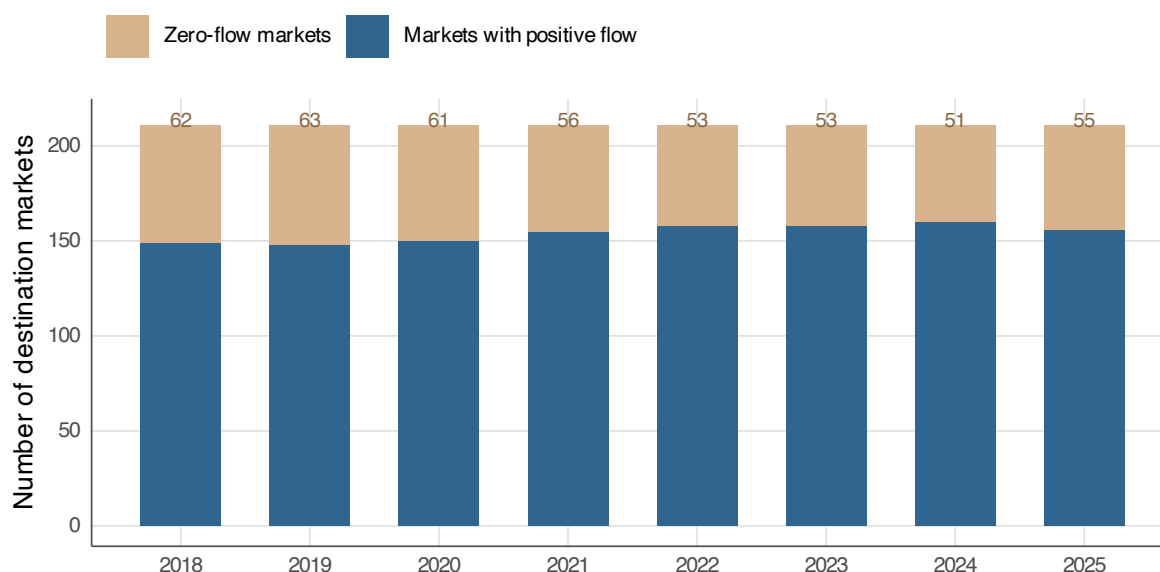


Figure 2.6: Active markets and zero trade flows by year, 2018–2025

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15.

The distribution of zeros across countries is informative for the estimation strategy. Destination markets fall into three groups. One hundred and thirty-four recorded positive exports in every year and account for essentially all recorded trade, 11079 million euros of the 11084 million total. Thirty-eight were served intermittently, appearing in some years and not others, and account for 5.5 million euros between them. The remaining thirty-nine recorded no exports at all. The third group is not randomly distributed: thirteen are in Asia-Pacific, ten in sub-Saharan Africa, eight in the Middle East and North Africa, six in non-EU Europe and two in Latin America and the Caribbean. They are small, distant, or subject to restrictions on alcohol, and in several cases all three.

These observations are the ones a log-linear specification cannot accommodate. Their exclusion, which is the procedure adopted by the reference study in this literature through the restriction of the sample to the forty leading destinations, does not merely reduce the number of observations: it removes from the sample precisely those markets about which the question of untapped export potential is being asked. The methodological argument advanced by Silva and Tenreyro (2006), which the following chapter takes up in operational terms, is that the Poisson pseudo-maximum-likelihood estimator accommodates these observations directly in the multiplicative form of the gravity equation, without transformation and without arbitrary additive constants.

The case for that estimator does not rest on the zeros alone. The second of Santos Silva and Tenreyro's two arguments, the inconsistency of log-linear estimation under heteroskedasticity, applies independently of whether any zero flows are present, and Chapter 3 examines both. The zeros matter for a different and complementary reason: without them the question of export potential in unserved markets cannot be posed at all.

The never-served group is one the sample design has to create rather than observe. A panel assembled from trade records alone cannot contain it: a country that has never appeared in a customs declaration leaves no row to extract. Constructing the frame from an independent universe of economies, and treating the absence of a record as a zero rather than as a missing value, is what makes these markets visible at all; in an earlier vintage of this dataset, built

directly from the flows, the group was empty. Tariff protection also differs systematically across the three groups: markets served throughout face an average applied rate of 9.4 per cent, markets never served 13.5 per cent, and markets served intermittently 16.9 per cent. That the intermittent group is the most protected, rather than the never-served one, suggests that tariffs in this sample do not so much foreclose access as make it precarious. These seventy-seven countries constitute the population on which the rotational validation design described in Chapter 4 operates: by placing each zero-flow market in turn in the test sample, the model is required to produce a prediction for markets it has not observed in the estimation, and the resulting figure is a genuine out-of-sample forecast rather than an in-sample residual. This design resolves the circularity objection to which residual-based measures of export potential are exposed, and it is a further reason the zeros are retained here rather than discarded.

Figure 2.7 shows the full distribution behind these averages, and it qualifies them. Among markets served throughout, the median applied tariff is zero: the group mean of 9.2 per cent is driven by a small number of highly protected destinations, India at 150 per cent and Indonesia at 90 per cent among them, which the appellation reaches despite the barrier. The intermittent group has no such tail, but its median is twenty per cent, so the higher mean reflects a broadly protected group rather than a few extreme cases. Markets never served sit between the two, with a median close to zero: for many of them the obstacle is not the tariff at all, but size, distance or the legal availability of alcohol.

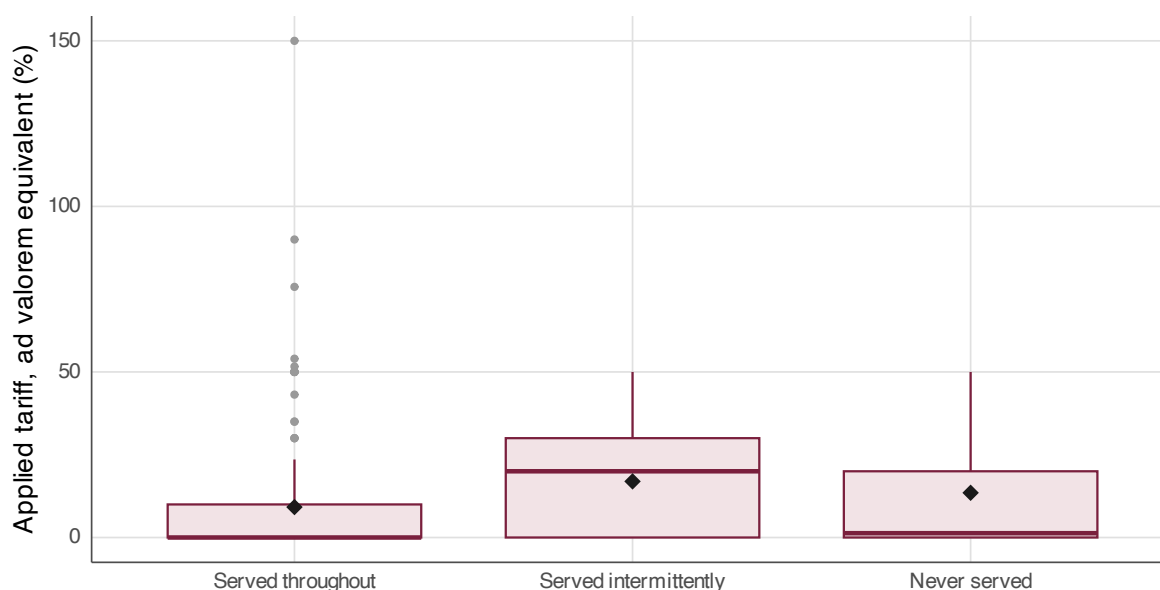


Figure 2.7: Applied tariff by destination group, 2025

Source: own elaboration on ITC Market Access Map and ISTAT Coeweb data. Diamonds mark group means; ad valorem equivalents for specific duties are computed on Prosecco unit values.

2.7 Descriptive statistics of the model variables

Table 2.6 reports the descriptive statistics of the variables entering the gravity specification, computed on the final panel of 1688 observations.

Table 2.6: Descriptive statistics, 2018–2025 (N = 1688)

Variable	Obs.	Mean	Std. dev.	Minimum	Maximum
Export value (€)	1688	6566533	36851000	0	490666667
Export quantity (kg)	1688	1585241	9123600	0	107106449
Unit value (€/kg)	1234	5.54	5.50	1.00	175.29
Partner GDP (USD bn)	1633	478.9	2217.0	0.05	30769.7
Partner population (mn)	1688	37.2	142.7	0.009	1463.9
GDP per capita (USD)	1633	20170	29990	216	288000
Distance (km)	1688	6090	4244	230	18570
Contiguity (0/1)	1688	0.02	0.15	0	1
Landlocked (0/1)	1688	0.18	0.39	0	1
EU membership (0/1)	1688	0.13	0.33	0	1
Free trade agreement (0/1)	1688	0.39	0.49	0	1
Applied tariff (% AVE)	1688	11.5	20.3	0	150.0
Alcohol consumption (l pc)	1464	4.54	3.82	0.00	15.20
Domestic wine production (l pc)	1688	4.19	12.58	0.00	71.46

Source: own elaboration on ISTAT Coeweb (CN 2204.10.15), World Bank WDI, CEPII GeoDist, DG Trade, ITC Market Access Map, WHO Global Health Observatory and OIV data. Unit value is undefined for the 454 zero-flow observations. Domestic wine production is zero for 160 of the 211 destinations; the 51 producing countries absorb 84.2% of export value.

Two features of the distribution have direct methodological implications. The first is the extreme dispersion of the dependent variable: the standard deviation of export value exceeds its mean by a factor of five, and the ratio of maximum to the median of positive flows is of the order of two thousand. This is the canonical signature of heteroskedasticity in trade data, and it is the second of the two arguments advanced by Silva and Tenreyro (2006) against log-linear

estimation, independent of the question of zero flows. A log transformation of such a distribution does not merely rescale it: under heteroskedasticity the expectation of the log-transformed error depends on the regressors, and the resulting estimates are inconsistent.

The second is the incidence of the binary regressors. Contiguity takes value one for only 2% of observations, corresponding to the five bordering countries present in the sample — Austria, France, San Marino, Slovenia and Switzerland, which means the coefficient on this variable is identified by a very small number of cross-sectional units and should be interpreted with corresponding caution. European Union membership covers 13% of observations, free trade agreements 39%. The within-country variation of these two variables is thinner than their cross-sectional incidence suggests: only the United Kingdom changes European Union status during the sample period, and only nine partners enter into a new agreement. Whether this scarcity of within variation, rather than the size of the sample, is what limits the precision with which the institutional effects can be estimated is a question Chapter 3 takes up.

The extreme values of the unit value distribution, reaching 175.29 euros per kilogram, correspond to shipments of a few dozen kilograms to marginal destinations and are plausibly sample consignments or promotional shipments. Four observations exceed twenty euros per kilogram — Turkmenistan in 2024, Togo in 2025, Fiji in 2020 and New Caledonia in 2018 — with consignments of between 23 and 1,850 kilograms and a combined value of 75 thousand euros, or 0.0007% of total exports. Twenty euros per kilogram is adopted as the threshold above which an observation is treated as a price outlier: it lies well above the 99th percentile of 13.30 and isolates precisely those four cases. These observations do not affect the Poisson estimates, which weight observations by their level rather than by their ratio, but they are excluded from any statistic computed on average prices.

The tariff variable requires one further note, since it is the only regressor in the specification that could not be taken from a single published source. Applied rates were collected market by market from the ITC Market Access Map for the destinations that account for the bulk of trade, and extended to the remainder through a documented cascade: rates were inherited where a country belongs to a customs union or to a trade agreement whose terms were verified for another member, and only where no such inheritance was available was a macro-area median used. Every observation carries a tag recording which level applied. Section 2.8 documents the cascade in full and reports the resulting coverage. This asymmetry between coverage by observation and coverage by value is a structural feature of the panel rather than a limitation of the collection, and Chapter 3 reports estimates both on the full sample and restricted to markets with a verified rate.

2.8 The construction of the tariff variable

Of the variables described in this chapter, the applied tariff is the only one not available as a ready series. Distances are published by CEPII, incomes by the World Bank, alcohol consumption by the World Health Organization; the duty that each of 211 destinations levies on Italian Prosecco is not published anywhere in a form usable for a gravity model, and had to be assembled. This section documents how, since the resulting figures carry a weight in Chapter 3 that would not be justified without an account of their provenance.

Three difficulties arise. The rate that applies is not necessarily the most-favoured-nation rate, since many destinations grant the European Union preferential access; a substantial minority of

countries levy duties as a fixed amount per unit of volume or weight rather than as a percentage of value, and these must be converted before they can enter a model; and no source covers every destination, so the residual has to be imputed by a rule stated in advance rather than case by case.

On the first point the collection follows a fixed priority of sources: the Market Access Map of the International Trade Centre, then the WITS-TRAINS database of the World Bank, then the Access2Markets portal of the European Commission. Where a preferential rate exists for European Union origin it is the one recorded, since that is the rate Italian exporters face. Of the 211 destinations, 109 apply a preferential rate and 102 the most-favoured-nation rate. The duty as displayed at source is transcribed verbatim alongside the numerical value, together with the date of consultation, so that any figure can be traced back to what the source stated on 21 August 2026.

The second point required the greater part of the work. Eleven destinations levy a specific duty — a fixed amount per litre in ten cases and per hundred kilograms of gross weight in the remaining one — and such a duty cannot enter a gravity equation until it is expressed as a share of value. The conversion depends on the price of the product, and this is where the published figures could not be used. Market Access Map computes *ad valorem* equivalents on the average unit value of the six-digit heading HS 220410, a category in which Champagne weighs heavily; since Champagne sells at several times the price of Prosecco, dividing the same fixed duty by that inflated reference price makes the burden appear far lighter than it is. The equivalents used here are therefore recomputed on the unit value of Prosecco itself, taken from the export data and computed year by year rather than once, since the unit value moves over the period. The price used is that of the flow being taxed, which is the accurate conversion in principle, since the duty is levied on those shipments and not on a sector average. It carries a cost that should be stated: the unit value of a flow is its value divided by its mass, so for the eight destinations whose applied duty is specific and whose own price was available the explanatory variable contains the dependent variable in its denominator. Section 3.6 re-estimates the model with the conversion performed on the aggregate Italian unit value of the year, a price common to all destinations and therefore independent of any one of them, and reports the consequences.

The United States illustrate the difference. The most-favoured-nation duty on sparkling wine is 19.8 cents per litre. Converted at the Prosecco unit value of 4.60 euros per kilogram, this amounts to 17.73 euros per hundred kilograms net, or 3.86% of value; computed on the category average of the six-digit heading it would be roughly half as much. The conversion requires two further parameters, both recorded in the collection workbook: a density of 0.99 kilograms per litre and a ratio of gross to net mass of 2.2 for bottled sparkling wine, which accounts for the weight of the glass, the closure, the label and the share of the carton. Exchange rates are the annual averages published by the European Central Bank.

The third point concerns coverage. The destinations were ordered by their share of 2025 export value and the first forty, which account for over 99% of it, were verified individually at source. Below that threshold individual verification would have consumed effort disproportionate to the value at stake and a cascade of rules was applied instead, in a fixed order: a member state of the European Union receives the internal rate of zero; a member of a customs union receives the common external tariff of that union; a dependent territory receives the rate of the jurisdiction whose customs territory it belongs to; and a destination for which none of these applies receives the rate of a verified country in the same region, or failing that the regional median.

Table 2.7: Provenance of the tariff variable

Origin of the rate	Dest.	Obs.	Share obs.	Share value
Verified at source	48	384	22.7%	68.1%
European Union internal rate	29	232	13.7%	30.7%
Verified country of the same region	52	416	24.6%	0.6%
Customs union common external tariff	24	192	11.4%	0.1%
Rate of the parent customs territory	18	144	8.5%	0.1%
Regional median	40	320	19.0%	0.3%
Total	211	1688	100%	100%

Destinations counts the 211 partners of the panel; observations counts the 1688 country-year cells, since the origin of the rate does not vary over time. Shares are computed on the observations and on export value respectively. Source: own elaboration; rates consulted on 21 August 2026 from ITC Market Access Map, WITS-TRAINS and the European Commission Access2Markets portal.

The distribution of effort follows the distribution of value. The rates verified individually at source, together with those that follow directly from European Union membership, cover 36.5% of the observations but 98.8% of the export value; adding the rates inherited from a customs union or a parent territory brings coverage to 81.0% of observations and 99.7% of value. The residual imputed by regional rule concerns 19.0% of the observations and three tenths of one per cent of the value, and consists almost entirely of small destinations that import little or nothing. An error in that residual could not materially affect any estimate in this thesis.

Two limitations should be recorded. The gross-to-net ratio of 2.2 rests on an assumed glass weight of approximately 800 grams for a 0.75 litre bottle, which is a plausible figure for sparkling wine but not one taken from a supplier specification; it affects a single destination, the only one levying a duty on gross weight. And the collection records the rate in force at the date of consultation, applied backwards over the period except where a change is known and documented; Section 3.7 shows that only twelve destinations record a change in the applied rate within the panel, and discusses what follows from that for the identification of the tariff coefficient.

2.9 Summary

This chapter has assembled the dataset on which the analysis rests. Italian exports of Prosecco, identified by the combined nomenclature code 2204.10.15, were extracted from ISTAT Coeweb for the years 2018 to 2025 and matched to the universe of economies recognised by the World Bank rather than to the set of destinations Italy has actually supplied. The resulting panel is balanced at 1688 country-year observations across 211 partners, and includes by construction the 454 cells in which no trade took place.

That construction is the principal methodological choice of the chapter and it follows from the question the thesis poses. A dataset built on observed flows can describe the markets Prosecco reaches but cannot say anything about those it does not and the 39 destinations that record no imports in any year would simply be absent from it. Retaining them costs nothing in terms of value — they contribute none — but it determines what can be asked in Chapter 4 and it requires the estimator discussed in Chapter 3.

Two errors in the source extraction were identified and corrected. The series recorded under the name of Russia in fact contained Belarusian values, while Russia itself, the fifth market by value in 2025, and Thailand were absent altogether; both omissions arose from the Italian country names used by the extraction interface, and both were detected by cross-checking against the mirror statistics of UN Comtrade. Correcting them raises annual totals by approximately five per cent and alters every concentration measure reported below. A second extraction at monthly frequency, described in Section 2.2, reproduces the corrected annual series exactly for all 126 overlapping country-year cells, which provides an independent check on the panel.

The descriptive evidence establishes three facts that Chapter 3 takes as its point of departure. Export value grew by 93.6% over the period, from 916 to 1772 million euros, with the pandemic year registering a fall in value of 4.1% against an increase in quantity of 2.2%. Concentration declined in every single year of the panel, the Herfindahl index falling from 1953 to 1320 and the combined share of the United States and the United Kingdom from 59.1% to 47.7%. And the extensive margin proved remarkably stable: of the 211 destinations, 134 recorded a positive flow in all eight years, 38 traded intermittently and 39 never at all, with the annual count of zero-flow markets varying only between 51 and 63. Chapter 3 uses these data to estimate the determinants of the flows described here and Chapter 4 uses the resulting model to ask what the absence of flow in the remaining markets implies.

Chapter 3

Estimating the gravity model

3.1 The empirical model

The gravity equation relates bilateral trade to the economic size of the trading partners and to the frictions that separate them. In the single-exporter setting adopted here, where the origin is fixed and only the destination varies, the exporter's own characteristics are common to every observation within a year and are absorbed by the time dimension. What remains to be explained is the distribution of Italian Prosecco exports across destinations and over time.

The specification estimated in this chapter takes the multiplicative form:

$$\begin{aligned} X_{(j,t)} = \exp \big(& \beta_0 + \beta_1 \ln \text{GDP}_{(j,t)} + \beta_2 \ln \text{POP}_{(j,t)} + \beta_3 \ln \text{DIST}_{(j)} \\ & + \beta_4 \text{CONTIG}_{(j)} + \beta_5 \text{LANDLOCKED}_{(j)} + \beta_6 \text{EU}_{(j,t)} \\ & + \beta_7 \text{FTA}_{(j,t)} + \beta_8 \text{ALCOHOL}_{(j)} + \beta_9 \ln(1 + \tau_{(j,t)}) \\ & + \beta_{10} \text{WINE}_{(j)} + \delta_{(t)} \big) \cdot \eta_{(j,t)} \end{aligned} \quad (3.1)$$

where $X_{(j,t)}$ denotes Italian exports of Prosecco to destination j in year t , measured in euros; GDP and POP are the destination's gross domestic product and population; DIST is the population-weighted bilateral distance from CEPII; CONTIG and LANDLOCKED are indicators for a shared land border with Italy and for the absence of sea access; EU and FTA indicate membership of the European single market and the existence of a preferential trade agreement with the Union; ALCOHOL is per capita alcohol consumption; τ is the applied ad valorem equivalent tariff on the product; WINE is domestic wine production per head; and $\delta_{(t)}$ is a full set of year fixed effects. The multiplicative error $\eta_{(j,t)}$ is assumed to satisfy $E[\eta \mid \cdot] = 1$, which is the only distributional assumption the estimator of Section 3.2 requires.

Two features of this specification depart from the canonical formulation and should be stated at the outset. The first concerns multilateral resistance. Anderson and Van Wincoop (2003) showed that bilateral trade depends not only on the frictions between two countries but on the trade costs each faces with all its other partners, and that omitting these terms biases the estimated effect of bilateral barriers. The standard remedy is to include exporter and importer fixed effects. That remedy is unavailable here: with a single exporter, destination fixed effects would be collinear with every time-invariant destination characteristic and would absorb dis-

tance, contiguity and the landlocked indicator, which are the trade-cost variables of interest. The specifications reported below are therefore of the single-exporter reduced form standard in the product-level literature, and the trade-cost coefficients should be read as incorporating whatever correlation exists between bilateral distance and the general remoteness of each destination.

The second concerns the demand-side and supply-side controls, which are not part of the canonical gravity equation and enter here for reasons specific to the product. Prosecco is a differentiated consumer good with a strong income gradient, and its market is shaped both by the propensity to consume alcohol at all — which varies enormously across destinations for cultural and legal reasons — and by the presence of a domestic sparkling wine industry that supplies a close substitute. Section 3.3 shows that the first of these is not optional: without it the model fails a specification test.

The domestic production variable requires a word on its construction. Output figures are drawn from the International Organisation of Vine and Wine. For the twenty countries producing more than one million hectolitres, which the OIV reports individually, the source is Table 2 of the State of the World Wine Sector in 2025 (OIV, 2026), which gives annual series for 2020 to 2025; for a further nine producers below that threshold but relevant to the analysis — Japan, Switzerland, Bulgaria, Croatia, Ukraine, the Czech Republic, Slovenia, Serbia, Israel and Albania — the figures were extracted from the OIV statistical database in August 2026. Twenty-two smaller producers, none of which appears among the markets discussed in Chapter 4, retain approximate values compiled from OIV annual reports. Output was converted to litres per inhabitant using World Bank population figures and enters the specification in tens of litres per head. Italy is absent from the variable, being the exporter rather than a destination.

Two features of the variable follow from the way it is compiled. It does not vary over time: the published table covers 2020 to 2025 while the panel begins in 2018, and the period mean was used rather than a series with two missing years. And it is zero for 160 of the 211 destinations, which produce no wine at all — a distribution that turns out to matter for the functional form, as Section 3.3 shows.

A word is due on the treatment of the pandemic, which Chapter 1 identified as one of the two shocks shaping the period. The dataset contains indicators for the pandemic years and for the years that follow, but neither appears in the specification above, because both are exact functions of the calendar year and are therefore perfectly collinear with the year fixed effects. The fixed effects do the same work in a less restrictive way: rather than imposing that 2020 and 2021 share a common shift, they allow each year its own level, and the pandemic is absorbed along with the exchange rate movements, the freight disruptions and everything else common to all destinations within a year. The pandemic indicator reappears only in Section 3.5, where the benchmark specification being replicated uses a time trend and a Covid dummy in place of year fixed effects; there it takes a coefficient of 0.045 and is not statistically significant, as in the published study.

The estimation sample comprises 1633 of the 1688 country-year observations described in Chapter 2. The 55 excluded cells are those for which the World Bank reports no gross domestic product figure: they concern almost entirely the provisional year 2025 and a group of micro-territories, and together they account for 0.12% of export value over the period. The exclusion applies to the estimation only — the dataset itself remains complete at 1688 observations — and the descriptive statistics of Chapter 2 are computed on the full panel. Standard errors are clustered by destination throughout, since the residuals of a given market are correlated across

the eight years of the panel.

3.2 The choice of estimator

The gravity equation is multiplicative in its structural form, and the conventional route to estimation is to take logarithms of both sides, which renders it linear in the parameters and permits ordinary least squares. The transformation is so nearly universal in the applied literature that it is rarely defended explicitly; Head and Mayer (2014), in their survey of gravity estimation, treat the choice between the log-linear and the multiplicative specification as one of the central methodological decisions facing the applied researcher, and it is treated as such here. In the present setting the logarithmic route cannot be adopted without difficulty, for two reasons examined below: it requires discarding a quarter of the sample, and it introduces a bias that falls precisely on the quantity this thesis sets out to measure.

Table 3.1 compares three ways of estimating the same equation on export value, using for the moment the core gravity variables alone. The first column reports ordinary least squares on logarithms, necessarily restricted to the 1215 observations recording positive trade. The second retains the full sample by adding one euro to every flow before taking logarithms. The third is Poisson pseudo-maximum likelihood, estimated on all 1633 observations.

Table 3.1: Determinants of Prosecco export value under three estimators

	(1) OLS on logs positive flows	(2) OLS on logs (value + 1)	(3) PPML full sample
ln GDP	1.326 ***	1.937 ***	1.736 ***
ln Population	−0.677 ***	−0.664 ***	−0.756 ***
ln Distance	−0.191	0.755 *	−1.008 ***
Contiguity	0.487	−0.436	−0.401
Landlocked	−0.291	0.491	−0.463
European Union	1.979 ***	6.370 ***	−0.352
Free trade agreement	0.646 **	3.460 ***	0.182
Observations	1215	1633	1633
of which zero flows	0	418	418
R² / squared correlation	0.732	0.527	0.777
RESET (p-value)	0.109	0.057	0.045

*Year fixed effects included in all columns; standard errors clustered by destination. For column (3) the goodness-of-fit measure reported is the squared correlation between observed and fitted values. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.*

Source: own elaboration on the dataset described in Chapter 2.

The most immediate difference between the columns is the size of the estimation sample. The logarithm of zero is undefined, so column (1) rests on 1215 observations rather than 1633. The 418 discarded cells contribute nothing to export value — that is what makes them zeros — and their exclusion might therefore appear inconsequential, but it is not. They constitute a quarter of the sample, and they are precisely the markets whose absence of trade Chapter 4 sets out to

explain. An estimator that cannot accommodate them cannot answer the question this thesis poses.

Column (2) illustrates the expedient most commonly adopted in response, which is to add a small constant to every flow so that the logarithm becomes defined everywhere. Its appeal is obvious: nothing is discarded, and the coefficients appear to retain their interpretation as elasticities. Head and Mayer (2014) advise against the practice on the ground that the results depend on the units of measurement and the interpretation of the coefficients as elasticities is lost, and Chen and Roth (2024) establish the point formally. When the outcome has a mass point at zero, the estimates obtained from log-like transformations depend on the units in which the outcome is measured, and by rescaling those units one can obtain an effect of essentially any magnitude. The reason is that the additive constant determines how much weight the transformation places on the movement from zero to a positive value, and that weight has no economic justification. In the present case a constant of one euro and a constant of one thousand euros are equally defensible choices that produce different answers.

The instability is visible in the estimates themselves. The distance coefficient in column (2) changes sign, becoming positive and marginally significant, which would imply that Prosecco travels more readily to remote markets than to nearby ones. The column is reported here only to establish that the device does not solve the problem.

The second difficulty is less visible in the coefficients and more consequential for what follows. A model estimated on logarithms delivers an estimate of the conditional expectation of the logarithm of exports, not of exports themselves. Recovering the latter requires exponentiation, and at this point Jensen's inequality intervenes: for any convex function the expectation of the transformed variable is at least the transformation of the expectation, with strict inequality whenever the variable is not degenerate. The exponential function is convex, so exponentiating an estimated mean of log exports necessarily understates the mean of exports itself, and the gap widens as the dispersion of the residuals grows. The transformation back to euros therefore introduces a systematic downward bias rather than an occasional one. Table 3.2 reports its size in the present data.

Table 3.2: Ability of each model to reproduce aggregate exports, 2018–2025

	Predicted total (€ mn)	Deviation
Observed value	11071.1	—
OLS on logs, naive exponentiation	4191.5	−62.1%
OLS on logs, Duan smearing correction	10592.1	−4.3%
PPML	11071.1	0.000%

Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15, extracted August 2026. The smearing factor is the non-parametric retransformation correction proposed by Duan (1983).

Exponentiating the fitted values of the log-linear model recovers 4191.5 million euros against the 11071.1 million actually exported. Duan (1983) proposed a correction for exactly this problem: rather than assuming a distribution for the error, the exponentiated predictions are scaled by the average of the exponentiated residuals, which is a consistent estimate of the missing factor provided the errors are identically distributed. Applying it reduces the shortfall to 4.3%,

a substantial improvement bought at the price of an assumption — identical distribution across observations — that the heteroskedasticity discussed below makes doubtful.

The comparison with the last row of the table should be read with care, and it is not the argument for the estimator adopted here. Poisson pseudo-maximum likelihood reproduces the aggregate exactly, but it does so by construction: the first-order condition associated with the constant term sets the sum of the residuals to zero, so the sum of fitted values equals the sum of observations whatever the quality of the fit. The log-linear model satisfies the corresponding condition in logarithms, where its own aggregate is matched exactly; that it fails to match the aggregate in levels is a consequence of Jensen's inequality rather than evidence of a defect. Judging the two estimators on their ability to reproduce a total in euros is therefore to judge them on ground that belongs to one of them.

What the table does establish is a practical point. Any use of the log-linear model to produce predictions in euros requires a retransformation step, that step requires a correction, and the correction rests on an assumption the data do not support. The Poisson estimator requires no such step. This is a convenience rather than a demonstration of superiority, and the demonstration is deferred to Chapter 4, where the two estimators are compared on their ability to predict destinations withheld from estimation. Out-of-sample performance is not guaranteed by any first-order condition, and is therefore the appropriate ground on which to choose between them.

The two considerations above concern what the log-linear model cannot do. A third concerns whether it is consistent at all, and it is the argument on which the modern literature principally rests. Silva and Tenreyro (2006) observe that taking logarithms transforms the error multiplicatively: if exports equal their conditional mean times a disturbance, the logarithm of exports equals the logarithm of the mean plus the logarithm of the disturbance. Ordinary least squares on that equation is consistent only if the expectation of the log disturbance is independent of the regressors. That condition is stronger than it appears, because by Jensen's inequality the expectation of a log depends on the entire distribution of the disturbance and not on its mean alone. If the variance of exports depends on the regressors — as it does whenever large markets are more volatile in absolute terms than small ones — then the expectation of the log disturbance will generally depend on them too, and the estimates are inconsistent.

Whether that is the case in the present data is an empirical question, and the answer is unambiguous. The test of Breusch and Pagan (1979) regresses the squared residuals of the log-linear model on its regressors and asks whether they have any explanatory power: under homoskedasticity the number of observations times the resulting coefficient of determination is distributed as a chi-squared variable with as many degrees of freedom as there are regressors. Here the statistic is 139.7 on 14 degrees of freedom, corresponding to a p-value below 10^{-20} . Homoskedasticity is rejected, and with it the condition on which the consistency of the log-linear model depends.

It is possible to go further and characterise the form of the variance. Santos Silva and Tenreyro suggest testing whether the conditional variance is proportional to a power of the conditional mean, which nests the leading cases: an exponent of one corresponds to a variance proportional to the mean, as the Poisson distribution implies, while an exponent of two corresponds to a variance proportional to the square of the mean, which is the case in which the log-linear specification would be appropriate. The test regresses the logarithm of the squared Poisson residuals on the logarithm of the fitted values, and the slope estimates the exponent. Here it is 1.823 with a standard error of 0.032, and both unity and two are rejected. The estimated exponent lies closer to two than to one, and the point deserves to be made explicitly rather than

passed over, since it is the one respect in which the evidence is not unambiguously favourable to the estimator adopted here.

It does not, however, favour the log-linear model. The Poisson estimator is a pseudo-maximum likelihood estimator in the sense of Gourieroux et al. (1984): it does not assume that exports are Poisson-distributed, and indeed they plainly are not, being continuous and highly dispersed. What it requires is only that the conditional mean be correctly specified, because the estimating equations depend on the data solely through the difference between observed and fitted values. Consistency therefore survives any form of the variance function, at a cost in efficiency when that function departs from proportionality. The log-linear specification enjoys no such protection, since its consistency rests on a condition about the log error that the heteroskedasticity test has just rejected.

The three columns of Table 3.1 do not differ in detail alone, and the divergences follow a pattern that the preceding discussion helps to interpret. Estimation on logarithms recovers the effect of each regressor on the geometric mean of exports among markets that already import; Poisson pseudo-maximum likelihood recovers the effect on exports measured in euros, and therefore weights large markets far more heavily. The two quantities need not coincide, and here they do not. Distance is statistically indistinguishable from zero in column (1) and becomes a precisely estimated elasticity of approximately unity in column (3); the pattern is consistent with distant markets that do import doing so in small volumes, so that their presence weakens the log-linear estimate while their limited weight in euros allows the Poisson estimate to reflect the concentration of exports in nearby markets. The indicator for European Union membership moves from +1.979 to -0.352 and loses significance, a reversal returned to in Section 3.4.

The evidence taken together supports estimation by Poisson pseudo-maximum likelihood on the complete sample. The remainder of this chapter adopts that estimator, and the log-linear specification is retained only as a point of comparison. The arguments given here are of an *a priori* kind: one estimator cannot use a quarter of the observations, and the consistency of the other requires a condition the data reject. Neither concerns how accurately either model predicts. That question is taken up in Chapter 4, where the two are compared on destinations withheld from estimation, both because the theoretical case deserves an empirical counterpart and because the accuracy of the chosen model — not merely its superiority over the alternative — determines what the analysis of export potential can conclude.

3.3 Building the specification

Having settled the estimator, the specification itself has to be constructed. The procedure followed here adds the non-canonical regressors one at a time and checks, at each stage, whether the resulting conditional mean is correctly specified. The instrument for that check is the specification-error test of Ramsey (1969), in the form adapted to multiplicative models by Silva and Tenreyro (2006): the square of the linear index is added to the regressors and its significance tested; a rejection indicates that the exponential form does not adequately describe the relationship. Proceeding in this way means the final specification is the outcome of a stated procedure rather than of a search across alternatives.

Table 3.3: Construction of the specification, PPML on export value

	(1) core gravity	(2) + alcohol	(3) + tariff	(4) + wine output
ln GDP	1.736 ***	1.408 ***	1.422 ***	1.332 ***
ln Population	−0.756 ***	−0.645 ***	−0.652 ***	−0.567 ***
ln Distance	−1.008 ***	−0.911 ***	−1.018 ***	−0.928 ***
Contiguity	−0.401	−0.496	−0.299	0.555
Landlocked	−0.463	−0.739 ***	−0.451 *	−0.703 ***
European Union	−0.352	−1.373 ***	−1.656 ***	−1.560 ***
Free trade agreement	0.182	−0.085	−0.311	−0.411
Alcohol consumption p.c.		0.277 ***	0.214 ***	0.248 ***
ln (1 + tariff)			−2.824 **	−3.461 ***
Wine output per capita				−0.210 **
Squared correlation	0.777	0.829	0.841	0.849
RESET (p-value)	0.045	0.123	0.155	0.115

All columns estimated by PPML on 1633 observations with year fixed effects; standard errors clustered by destination. Columns (2) to (4) include an indicator for the 224 country-year cells without WHO alcohol consumption data, not reported. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Source: own elaboration.

The first step is the one that matters most for the credibility of what follows. The core gravity specification of column (1) fails the RESET test at the 5% level, with a p-value of 0.045. Adding per capita alcohol consumption removes the rejection, raising the p-value to 0.123, and lifts the squared correlation from 0.777 to 0.829. The variable therefore entered the specification because the model was misspecified without it, not because it improved the fit; the improvement in fit is a by-product. The distinction matters, because a specification assembled by adding regressors until the fit stops improving is not a specification that can be defended.

The economic content of the result is straightforward. Prosecco is an alcoholic beverage, and the propensity to consume alcohol varies across destinations by more than an order of magnitude for reasons — religious, legal and cultural — that income and distance do not capture. A gravity model that ignores this variation misattributes it to the variables it does contain, which is what the functional-form test detects.

The tariff enters in column (3) with a coefficient of −2.824 and adds a further point to the squared correlation. Domestic wine production enters in column (4) with a coefficient of −0.210, significant at the 5% level. Its inclusion also raises the tariff coefficient in absolute value from −2.824 to −3.461, which is intelligible: several of the most heavily protected destinations are also wine-producing countries, and before the production variable is included the tariff coefficient absorbs part of the substitution effect that properly belongs to domestic supply.

One aspect of the production variable deserves comment because the choice is not obvious. It enters in levels — litres per head — rather than in logarithms, and the logarithmic form is in

one respect the more attractive of the two, since it yields a higher squared correlation and a coefficient of -0.306 that is also significant. It was nevertheless rejected, because it fails the functional-form test at a p-value of 0.015 where the level specification passes at 0.115.

The two forms are not alternative parameterisations of the same relationship but different economic hypotheses. In levels, each additional litre of domestic output per head reduces imports by the same proportion, so that moving from zero to ten litres has the same effect as moving from sixty to seventy. In logarithms the effect is concentrated at the bottom of the range and flattens thereafter, so that what matters is a doubling of output rather than an increment to it. Which of the two describes substitution between domestic sparkling wine and Prosecco is an empirical question, and the distribution of the variable makes it a consequential one: 160 of the 211 destinations produce no wine at all, and the logarithmic transformation compresses the upper range so severely that the distance between zero and one litre per head becomes comparable to the distance between sixty and seventy.

The consequences are visible in the predictions. For the large producers the two forms agree: Spain, at 71.5 litres per head, is predicted to import 78% less under the level form and 73% less under the logarithmic one. They diverge at the bottom, and there the logarithmic form is difficult to defend. It attributes to Japan, which produces 0.7 litres per head, a substitution effect of 15%, where the level form implies 1%; and it attributes to Germany, at 10 litres per head, an effect of 52%, which is three quarters of what it assigns to France despite a level of production six times smaller. Domestic output of that order does not constitute meaningful competition for an imported sparkling wine, and a specification implying that it does is describing something other than substitution.

The episode illustrates why the functional-form test rather than the fit statistic governs these decisions. The squared correlation measures how close the fitted values lie to the observations within the estimation sample; the RESET test asks whether the shape of the relationship is right. A misspecified curve can track the data well over the observed range while implying implausible behaviour elsewhere, and since Chapter 4 uses the model to predict markets it has not seen, the shape matters more than the proximity. The same principle governed the inclusion of alcohol consumption earlier in this section; here it operates against the specification that fits better, which is the stronger test of whether the principle is being applied consistently.

3.4 Results

Table 3.4 reports the preferred specification estimated on export value and, as a consistency check, on export quantity in kilograms. The third column reproduces the same specification under the log-linear estimator, for comparison with the literature.

Table 3.4: Preferred specification

	PPML value	PPML quantity	OLS on logs value
ln GDP	1.332 *** (0.147)	1.314 *** (0.153)	1.211 *** (0.106)
ln Population	-0.567 *** (0.197)	-0.522 ** (0.209)	-0.538 *** (0.119)
ln Distance	-0.928 *** (0.200)	-1.033 *** (0.211)	-0.289 * (0.149)
Contiguity	0.555 (0.414)	0.630 (0.492)	0.447 (0.687)
Landlocked	-0.703 *** (0.230)	-0.897 *** (0.246)	-0.448 * (0.250)

Table 3.4: Preferred specification (*continued*)

	PPML value	PPML quantity	OLS on logs value
European Union	−1.560 *** (0.564)	−1.605 ** (0.632)	0.835 * (0.506)
Free trade agreement	−0.411 (0.356)	−0.441 (0.377)	0.320 (0.292)
Alcohol consumption p.c.	0.248 *** (0.071)	0.257 *** (0.080)	0.158 *** (0.045)
ln (1 + tariff)	−3.461 *** (1.267)	−3.854 *** (1.450)	−1.216 (0.770)
Wine output per capita	−0.210 ** (0.093)	−0.235 ** (0.114)	−0.064 (0.074)
Observations	1633	1633	1215
Squared correlation	0.849	0.798	0.760
RESET (p-value)	0.115	0.165	0.030

Standard errors clustered by destination in parentheses. Year fixed effects and the alcohol-missing indicator included but not reported. Wine output is measured in tens of litres per head. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Source: own elaboration on ISTAT Coeweb, World Bank WDI, CEPII GeoDist, DG Trade, ITC Market Access Map, WHO Global Health Observatory and OIV data.

The two Poisson columns pass the functional-form test and the log-linear column fails it, at a p-value of 0.030. Value and quantity yield coefficients that are close to one another throughout, which is reassuring: the two dependent variables are measured independently in the customs declarations, and agreement between them suggests the model is capturing something real rather than a feature of one measure.

The income elasticity is 1.332, so exports rise slightly more than proportionally with the size of the destination economy. The negative coefficient on population must be read jointly with it: at a given level of aggregate income, more populous countries import less Prosecco. Reparameterising in terms of income per head — an algebraically equivalent transformation, since log GDP is the sum of log GDP per capita and log population — yields an elasticity of 1.332 with respect to income per capita and of 0.765 with respect to population. Prosecco follows purchasing power more closely than demographic mass, which is what one expects of a differentiated consumer good in the middle of the price range.

Distance carries an elasticity of -0.928 , close to the unit value typical of the gravity literature and considerably larger in magnitude than the -0.289 of the log-linear column. The landlocked indicator is negative and significant at -0.703 , consistent with the additional handling costs of a fragile bottled product that cannot be delivered directly to a port. Contiguity is not significant, as expected: the variable is identified by five countries only.

Per capita alcohol consumption carries a coefficient of 0.248, so a destination consuming one additional litre of pure alcohol per head imports approximately 28% more Prosecco. Domestic wine production enters negatively at -0.210 per ten litres per head. The magnitude is substantial when applied to the largest producers: Spain, at 71.5 litres of wine produced per inhabitant, is predicted to import 78% less Prosecco than an otherwise identical country with no domestic production. Chapter 4 shows that omitting this variable places Spain, Portugal and Chile at the head of the ranking of under-served markets, which is not plausibly a finding about Prosecco.

The tariff coefficient of -3.461 implies that a ten percentage point difference in the applied duty is associated with export value lower by 28.1%, and export quantity lower by 30.7%.

Read as a trade elasticity, the magnitude is close to the central tendency of the literature: the meta-analysis of structural gravity estimates in Head and Mayer (2014) reports a median of 3.78, and Simonovska and Waugh (2014) obtain 4.14. Section 3.6 shows that this coefficient is the least stable in the specification, and Section 3.7 explains why.

The result that requires most comment is the European Union indicator, which is negative and significant at -1.560 on value and -1.605 on quantity. Taken at face value it would say that membership of the single market reduces Prosecco imports, which is implausible. Four considerations bear on its interpretation, and the third is the one the evidence supports.

The first is that the coefficient is a cross-sectional association. A single destination changes European Union status within the observation window, so the estimate is identified almost entirely by comparisons between countries rather than by any change in status; Section 3.6 shows that where within-country variation does exist, the sign is the expected one.

The second is that the coefficient is not stable. It falls by half when the two largest destinations are excluded, which indicates that a substantial part of it reflects the position of the United States and the United Kingdom rather than a general regularity.

A third consideration concerns what the coefficient can measure once the rest of the specification is in place. The model controls simultaneously for European Union membership, for the existence of a free trade agreement and for the applied tariff. Since member states face no duty on intra-Union trade, the tariff advantage of membership is already absorbed by the tariff variable, and the same is true of whatever preferential access the agreement indicator captures. What remains attached to the European Union indicator is a residual: whatever distinguishes member states from other destinations after income, population, distance, alcohol consumption, domestic production, tariffs and trade agreements have all been accounted for. A residual of this kind has no natural interpretation, and it should not be read as the effect of membership.

The last possibility, and the one initially most attractive, is that the comparison group is anomalous. The destinations most similar to the European member states in income, distance and alcohol consumption include the United Kingdom, Switzerland and Norway, all outside the Union and all exceptionally heavy purchasers: Switzerland imports 5.90 euros of Prosecco per inhabitant each year and the United Kingdom 5.08, against a median of 1.20 among the member states. The arithmetic of the sample points in the same direction. In 2025 the United States and the United Kingdom together absorbed 47.7% of Italian Prosecco exports, while the twenty-seven member states of the Union absorbed 33.5% between them. Conditional on income, population and distance, it is therefore unsurprising that member states should display lower average flows than non-members: the composition of the comparison group is sufficient on its own to generate a negative coefficient, without any economic mechanism being required. The indicator might therefore be measuring a shortfall relative to an unusually strong control group.

This conjecture was tested directly, and it holds only in part. Adding an indicator for the three countries moves the European Union coefficient from -1.560 to -1.057 , an attenuation of a third; but excluding them from the sample altogether moves it in the opposite direction, to -1.769 . The three exceptional buyers account for some of the coefficient, not for most of it.

What does account for it emerges when the sample is restricted to European destinations alone, so that member states are compared with their immediate neighbours rather than with the rest of the world. The coefficient then ceases to be negative: it takes a value of 0.956 and is no longer statistically distinguishable from zero. The restriction is costly in precision — 437

observations across 56 destinations, of which 26 are member states — and the confidence interval is correspondingly wide, running from -0.46 to 2.38 . No positive effect of membership can be inferred from it.

What can be inferred is narrower but sufficient for the present purpose: the strong negative association of the full sample does not survive the restriction, and it therefore reflects a comparison between Europe and the rest of the world rather than one between members and non-members. The indicator appears to be capturing something regional that the gravity variables do not measure — plausibly the density of distribution networks, the familiarity of the product, or the concentration of tourism from Prosecco-consuming countries — and attributing it to the institutional variable that happens to be correlated with region. No interpretation is therefore placed on its coefficient in what follows.

The variable is nevertheless retained, because its omission is not neutral. Estimating the same specification without it moves the distance elasticity from -0.928 to -0.508 and the tariff coefficient from -3.461 to -1.317 , changes of 45% and 62% respectively, and reverses the sign of the free trade agreement indicator; the squared correlation falls from 0.849 to 0.774. Whatever regional variation the indicator absorbs is evidently correlated with distance and with tariff protection, and removing it transfers that variation to coefficients that do carry an interpretation. Retaining an uninterpreted control is preferable to contaminating the coefficients that matter. Chapter 4 examines the sensitivity of the potential estimates to this choice.

3.5 Comparison with the benchmark study

The closest antecedent to the present work is Echevarria Garay et al. (2025), a preprint that models flows to the 40 leading destinations between 2017 and 2023 and simulates the effect of a United States tariff. Establishing how the present work relates to it serves two purposes: it provides an external check on the dataset assembled in Chapter 2, and it clarifies that the two studies address different questions rather than reaching different answers to the same one.

Their treatment of zero flows illustrates the difference. Zeros are neither dropped nor transformed in that study, because the sample is constructed so that they do not arise: each of the forty markets records a positive flow in every year of their window. The additive constant of 0.001 that appears in their specification is applied to the tariff variable, in order to permit its logarithmic transformation where the applied duty is zero, not to the trade flow. The design is entirely coherent with the question they pose, which concerns the determinants of flows towards markets where Prosecco is already established. It cannot address the question posed here, which concerns the markets where it is not.

Their specification was replicated on the present data, restricting the sample to the same forty markets and to the years that overlap with theirs, and estimating by ordinary least squares on the logarithm of quantity as they do. Table 3.5 sets the two sets of estimates side by side.

Table 3.5: Replication of Echevarria Garay et al. (2025)

	Echevarria Garay et al. 2017–2023, TradeMap	Present data 2018–2023, ISTAT
ln GDP of importer	0.644 ***	0.606 ***
ln Distance	−0.424 ***	−0.606 ***

Table 3.5: Replication of Echevarria Garay et al. (2025) (*continued*)

	Echevarria Garay et al. 2017–2023, TradeMap	Present data 2018–2023, ISTAT
In Tariff	−0.130 ***	−0.094
Specific duty (indicator)	1.394 ***	2.247 ***
Time trend	0.147 ***	0.139 ***
Covid (2020–2021)	0.051	0.045
R-squared	0.522 (adjusted)	0.515
Observations	280	240

*Dependent variable: logarithm of export quantity in kilograms. OLS without fixed effects, standard errors clustered by destination in the replication. *** $p < 0.01$. Benchmark column from Table 2 of Echevarria Garay et al. (2025).*

Source: own elaboration on ISTAT Coeweb and ITC Market Access Map data.

The replication succeeds. The income elasticity, the time trend and the statistical insignificance of the Covid indicator coincide almost exactly, and the explanatory power is effectively identical, notwithstanding that the trade data derive from different sources — Italian customs declarations here, mirrored import statistics there — and that the tariff variable is constructed differently, as explained in Section 2.8. The agreement is informative precisely because the two datasets are independent: it indicates that the discrepancies documented in Chapter 2, and the corrections applied there, have not distorted the series relative to the published record.

Two coefficients differ appreciably. Distance is more strongly negative in the present data, at -0.606 against -0.424 . The tariff coefficient is smaller in magnitude and loses significance, at -0.094 against -0.130 ; it recovers both magnitude and significance under the estimator adopted in this chapter, which is consistent with the tariff effect being concentrated in the larger markets that the Poisson estimator weights more heavily.

The relationship between the two studies is therefore one of complementarity. Echevarria Garay et al. estimate the determinants of established flows and simulate a tariff shock within that set of markets. The present work extends the sample to the universe of destinations, which requires an estimator capable of handling zeros, and uses the resulting model to identify markets whose imports fall short of what their fundamentals would predict. The methodological divergence follows from the difference in question.

Their tariff simulation was calibrated on the increase to 30% announced in July 2025. The measure eventually implemented was of 15%, applied from 1 August. The scenario therefore remains a counterfactual rather than a description of what occurred, which is a further reason for identifying the effect of the realised measure directly, as Section 3.8 does.

3.6 Robustness

The estimates of the previous section rest on a single specification estimated on the full sample, and three questions naturally arise about their stability: how much of the result is driven by the largest destinations, whether the institutional coefficients survive when identified from within-country variation alone and how sensitive the findings are to error in the domestic production

data. Each is examined in turn.

Poisson pseudo-maximum likelihood weights observations by their level, so the largest destinations exert considerable influence on the estimates. Over the period the United States and the United Kingdom together account for approximately half of Italian Prosecco exports, and it is necessary to establish how much of the result depends on them. Table 3.6 reports the preferred specification re-estimated with each of them excluded in turn and with both removed together.

Table 3.6: Sensitivity to the exclusion of the leading markets

Sample	ln GDP	ln Distance	European Union	ln (1 + tariff)
Full sample	1.332	−0.928	−1.560	−3.461
Excluding the United States	1.148	−0.948	−1.104	−3.191
Excluding the United Kingdom	1.318	−0.685	−1.684	−2.574
Excluding both	1.033	−0.586	−0.788	−1.328

PPML on export value, preferred specification, year fixed effects. Source: own elaboration.

The income elasticity is reasonably stable, falling from 1.332 to 1.033 when both markets are removed, and distance weakens rather more, from -0.928 to -0.586 . The European Union indicator halves, which is among the reasons for the interpretation offered in Section 3.4.

The tariff coefficient is the least robust of the four, falling from -3.461 to -1.328 , a reduction of more than half. This is not surprising once the source of identification is understood: as Section 3.7 sets out, the applied tariff varies over time for twelve destinations only, and the United States is both one of them and the largest market in the sample, so removing it removes much of the variation on which the coefficient rests. The magnitude estimated here should accordingly be regarded as indicative rather than precisely determined, and the analysis of the 2025 United States measure in Section 3.8 provides the sharper evidence.

A separate concern attaches to the way the tariff variable is constructed rather than to the variation that identifies it. As Section 2.8 records, the ad valorem equivalents of specific duties were computed on the unit value of the flow being taxed, and for the eight destinations concerned that unit value is the flow's own value divided by its own mass. The dependent variable therefore appears in the denominator of the regressor, and the mechanical correlation runs in the same direction as the estimated effect: a flow of higher value carries a higher unit value, a lower equivalent and a smaller logarithm, so part of the negative coefficient may be arithmetic rather than economic. The concern is not negligible in scale, since the destinations affected account for about three tenths of export value and include the United States and Switzerland.

Recomputing the equivalents on the aggregate Italian unit value of each year, a price common to every destination and therefore independent of any one of them, allows the size of the effect to be assessed. The tariff coefficient moves from -3.461 to -3.117 , a reduction of a tenth and well within one standard error, and remains significant at the one per cent level; no other coefficient changes in the second decimal. The exercise in Chapter 4 is unaffected, since the eleven under-served markets and the aggregate gap of 142 million euros a year are identical under the two constructions. The mechanical component is thus present, as the sign of the change confirms, but small enough to leave both the estimate and the conclusions that rest on

it intact.

The institutional variables can be identified from within-country variation alone by adding a full set of destination fixed effects. The exercise is diagnostic rather than a candidate specification, since the fixed effects absorb distance, contiguity, the landlocked indicator, alcohol consumption and domestic wine production, all of which are constant within a destination, and the thirty-nine markets that never import are dropped for lack of variation, leaving 1354 observations on 172 destinations. Under this specification the European Union indicator becomes positive and significant, at 0.506 with a standard error of 0.094, implying export levels approximately 40% below the counterfactual following withdrawal. The estimate is identified by the United Kingdom, the only destination to change status within the window, and it cannot be separated from anything else specific to the British market after 2020; it is reported as suggestive rather than as a measurement of the effect of Brexit. The free trade agreement indicator remains insignificant at 0.086. The contrast with the cross-sectional estimate is nonetheless instructive: where within-country variation exists the European Union effect has the expected sign, and where it does not the cross-sectional coefficient reflects composition rather than membership, which is the reading advanced in Section 3.4.

The choice of functional form for the domestic production control was itself the outcome of a specification test, discussed in Section 3.3, and is not revisited here. A final concern is measurement error in the production data. These combine the published OIV country table, which covers producers above one million hectolitres, with database extractions for the smaller producers, as documented in Section 3.1. An earlier version of the analysis used approximate figures compiled from OIV annual reports, which for individual countries departed from the official series considerably: the Serbian figure was overstated by two thirds. Replacing them with the official data moved the production coefficient from -0.215 to -0.210 and left the ranking of markets in Chapter 4 unchanged. The control therefore operates through the large producers, where the figures were reliable in both versions, and the conclusions are robust to error among the smaller ones.

3.7 Limits of identification

Four features of the data constrain what these estimates can establish. Each is a property of the available variation rather than a defect of the estimation, and each bears on how the coefficients should be read.

Table 3.7: Within-country variation of the regressors

Variable	Destinations varying over	
	time	Out of
Alcohol consumption per capita	0	211
Domestic wine production	0	211
Bilateral distance	0	211
European Union membership	1	211
Free trade agreement	9	211
Applied tariff	12	211

Number of destinations for which the variable takes more than one distinct value across the eight years of the panel. Source: own elaboration.

The applied tariff varies over time for twelve destinations only, although these account for 30.9% of export value. The estimated elasticity is therefore largely cross-sectional in origin: it arises from comparisons between countries with differing levels of protection rather than from observed changes in commercial policy within a given market, and it remains exposed to whatever distinguishes protected from open destinations and is absent from the specification. The instability documented in Section 3.6 is the visible consequence. Section 3.8 addresses the question directly.

Per capita alcohol consumption is a single World Health Organization figure per country, repeated across all eight years, and domestic wine production is time-invariant for the reasons set out in Section 3.1. Both vary between destinations but never within one. They are therefore estimable in the cross-section alone and are dropped from any specification including destination fixed effects. This is not a defect of the data but a boundary on what can be identified, and it happens to fall in a convenient place: the cross-section is the dimension relevant to the estimation of export potential in Chapter 4, which asks how much a given market would absorb rather than how a given market responds to change.

The institutional variables carry very little variation over time. One destination changes European Union status within the window and nine enter a free trade agreement, so neither coefficient can be read as an estimate of the effect of accession or of agreement entry; both are cross-sectional associations, and Section 3.4 shows what that implies in the case of the European Union indicator. It should be noted that the free trade agreement variable does not lack variation altogether: 77 destinations hold an agreement throughout the period and 125 hold none, so the coefficient is identified from a substantial cross-section. Its consistent insignificance under the Poisson estimator is therefore a result rather than an artefact of the data, and one that is intelligible: the destinations covered by an agreement account for 27.3% of export value, a share dominated by markets whose imports the gravity variables already explain well, so that the agreement indicator adds little once income, distance and tariffs are in the specification.

Finally, and as set out in Section 3.1, the single-exporter design precludes the inclusion of multilateral resistance terms alongside the trade-cost variables of interest. The distance elasticity reported here therefore incorporates any correlation between bilateral distance from Italy and the general remoteness of each destination from world markets, and is not directly comparable with estimates from multi-exporter specifications that include importer fixed effects.

3.8 The United States tariff of 2025

Section 3.7 identified the tariff coefficient as the least securely estimated in the specification, on the ground that the applied rate varies over time for twelve destinations only. The United States tariff of 2025 offers the opportunity to escape that limitation. It is a dated policy change affecting one market and not the others, and the tariff coefficient can therefore be approached through a comparative design built around that change rather than inferred from a cross-section.

Annual data cannot serve this purpose. The measure entered into force on 1 August 2025, so the observation for that year blends seven months without the tariff and five with it, and it remains a single post-treatment observation in any case. The analysis in this section therefore uses a

second extraction from ISTAT Coeweb at monthly frequency, covering the twenty-one leading destinations from January 2020 to May 2026, for 1617 country-month observations. Summing the monthly figures reproduces the annual series of Chapter 2 exactly, to the euro, for all 126 country-year cells that overlap; the two extractions are independent, so this is also an external check on the panel.

The chronology of the measure is more involved than a single break date, and the design depends on getting it right. Tariffs on European goods were threatened from early 2025 and a baseline duty of 10% was in force from the spring. The political agreement was announced on 27 July, the executive order signed on 31 July and the rate of 15% applied from 1 August. Goods already loaded before that date, however, remained exempt provided they arrived before 5 October, so August and September are hybrid months in which shipments at the old and new rates coexisted. Anticipation matters at the other end of the window: United States imports of Italian sparkling wine rose sharply in November 2024, immediately after the presidential election, as importers filled their pipelines ahead of the expected measure.

The specification accordingly divides the period into five phases rather than two. The months to October 2024 form the baseline. The five months from November 2024 to March 2025 capture the post-election anticipation. The four months from April to July 2025 are those of the 10% baseline duty. August and September 2025 are the transition, and the eight months from October 2025 to May 2026 are the period of full treatment.

The estimator is Poisson pseudo-maximum likelihood with destination and month fixed effects, the month effects absorbing everything common to all destinations within a given month: seasonality, world demand, freight costs and exchange rate movements shared across markets. The coefficients on the interaction between the United States indicator and each phase measure how far the United States departs from the control markets in that phase. The composition of the control group determines what the comparison measures and is therefore set out in full. The pool consists of the twenty largest destinations for Italian Prosecco after the United States, which are the markets for which monthly customs data are available at the required level of detail and whose flows are large enough for a monthly series to be informative rather than erratic. The choice of a size criterion, rather than a similarity criterion applied market by market, is deliberate: it is fixed in advance of the analysis and cannot be adjusted to produce a preferred result. Three of the twenty were then removed, each for a reason established independently of Prosecco and dated before the measure under study. The United Kingdom left the single market in January 2021 and its trade with Italy has followed a distinct path since. Russia and Ukraine have been subject to war and to sanctions since February 2022, which disrupts both demand and the recording of trade. A market undergoing a shock of its own cannot stand for what the United States would have done in the absence of the tariff. Seventeen controls remain, spanning the European Union, the remaining European markets, Asia and Oceania, and the estimation sample comprises 1386 observations. The identifying assumption is not that these markets resemble the United States in level, which they do not, but that their exports would have moved in parallel with American exports over the window considered; Section 3.8 examines that assumption directly.

Table 3.8: Italian Prosecco exports to the United States relative to control markets

Phase	Months	Value	Quantity	Implied unit value
Anticipation (Nov 2024 – Mar 2025)	5	+19.7% ***	+13.1% **	+5.9%
10% baseline duty (Apr – Jul 2025)	4	–13.6% *	–13.0%	–0.7%
Transition (Aug – Sep 2025)	2	–4.0%	–5.3%	+1.3%
Full 15% tariff (Oct 2025 – May 2026)	8	+4.7%	+5.8%	–1.0%

PPML with destination and month fixed effects, 1386 observations on eighteen markets, January 2020 to May 2026; the months to October 2024 are the reference. Standard errors clustered by destination. Effects reported as $\exp(\beta)-1$. The implied unit value is the difference between the value and quantity coefficients. *** $p<0.01$, ** $p<0.05$, * $p<0.10$. Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15, extracted August 2026.

The clearest result in the table is the one that does not concern the tariff. In the five months following the presidential election, exports to the United States ran 19.7% above the control markets in value and 13.1% in quantity, and both estimates are statistically significant. The pattern reverses in the following four months, when the baseline duty was in force and shipments fell 13.6% below the controls. This is the signature of stockpiling: importers accelerated purchases in anticipation of a measure that had been announced but not yet imposed and then drew down the inventory they had accumulated. It is not an effect of the tariff and a specification that ignored the anticipation would have attributed the subsequent trough to the measure.

In the period of full treatment, from October 2025 to May 2026, exports to the United States stand 4.7% above the control markets in value and 5.8% above in quantity. Neither estimate is statistically distinguishable from zero. The point estimates are positive rather than negative; taken literally they would say that the tariff was followed by a slight increase, which is not a claim worth making, but the direction is worth noting because it is the opposite of what a substantial contraction would look like.

The implied unit value is flat throughout, moving by no more than one percentage point in either direction outside the anticipation window. The price at which Prosecco leaves Italy did not change. This confirms the descriptive finding of Chapter 2 and is, on reflection, what should have been expected: the duty is levied on the importer at the United States border, and there is no mechanism by which it would alter the euro price agreed between the Italian producer and its counterparty.

Figure 3.1 traces the same comparison month by month, taking the month before entry into force as the reference.

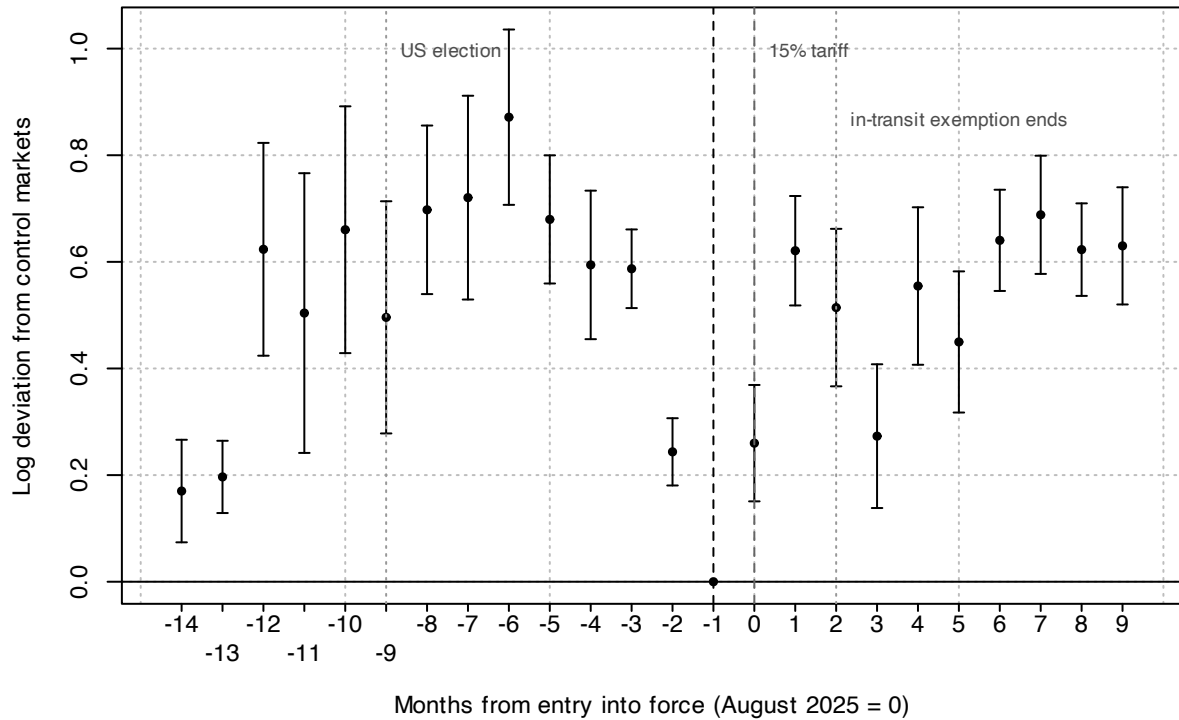


Figure 3.1: Italian Prosecco exports to the United States, deviation from control markets
Italian Prosecco exports to the United States relative to seventeen control markets, monthly, January 2020 to May 2026. Coefficients from a Poisson event-study specification with destination and month fixed effects; the month preceding entry into force is the reference. Dashed line: entry into force of the 15% tariff, 1 August 2025. Dotted lines: the United States election of November 2024 and the expiry of the in-transit exemption on 5 October 2025. The confidence bands are indicative only, since the number of interaction terms exceeds the number of clusters. Source: own elaboration on ISTAT Coeweb data, CN 2204.10.15.

Two checks bear on how much weight these estimates can carry, and neither is entirely reassuring. The first concerns the assumption that the United States and the control markets would have moved together in the absence of the measure. This cannot be verified directly, but a differential trend can be estimated over the baseline period; it comes to -0.20% per month with a standard error of 0.17% , which is not statistically significant but compounds to -10.9% over the fifty-eight months of the baseline. The parallel-trends assumption survives the test without being confirmed by it.

The second concerns inference. With a single treated unit, clustered standard errors are unreliable, and the conventional interval for the full-treatment coefficient — from -8.5% to $+19.8\%$ — is almost certainly too narrow. A permutation test provides a more honest assessment: estimating the same specification seventeen times, each time treating one of the control markets as though it had been subject to the tariff, yields a distribution of placebo coefficients with a standard deviation of 0.274 . The coefficient estimated for the United States is smaller in absolute value than eleven of the seventeen placebos, giving a p-value of 0.78 . The observed effect is indistinguishable from what would be obtained by designating an arbitrary market as treated.

These two results should be read together rather than separately. The conventional interval indicates that a severe contraction did not occur: had exports to the United States fallen by a quarter, the estimate would not have been what it is. The permutation test indicates that a moderate contraction would not have been detected, because the month-to-month volatility of individual markets is of the same order as the effect being sought. The honest statement is

therefore that Italian Prosecco exports to the United States show no sign of having declined in the eight months following the imposition of the tariff and that the design has power to exclude large effects but not small ones.

One limitation of interpretation deserves emphasis, because it constrains what can be concluded from any analysis of export data. These figures record what Italy shipped and at what price it was invoiced. They do not record what happened to the duty. A tariff of 15% on the customs value must be absorbed somewhere along the chain — by the importer, the distributor, the retailer or the final consumer — and the export statistics are silent on which. That volumes did not fall is consistent with the duty having been absorbed in distribution margins and equally consistent with its having been passed to consumers who did not reduce their purchases. Distinguishing between these would require United States retail prices, which lie outside the scope of this thesis.

Finally, the result bears on the benchmark study discussed in Section 3.5. Echevarria Garay et al. (2025) simulated the consequences of the increase to 30% announced in July 2025, and projected a substantial contraction of Italian sparkling wine exports. The measure eventually implemented was of 15% and at eight months no contraction is visible in the flows. The two findings are not in conflict: theirs is a counterfactual for a policy that was not adopted, and a simulation calibrated on a cross-sectional elasticity cannot anticipate how a particular market will absorb a particular measure. The comparison does, however, illustrate the value of the distinction drawn in Section 3.7 between an elasticity estimated across countries and the effect of an identified policy change.

3.9 Summary

This chapter has established the empirical model on which the analysis of export potential rests. The choice of Poisson pseudo-maximum likelihood over the conventional log-linear estimator was made on two grounds that are established here and a third that is established in Chapter 4. The first is that it retains the 418 zero-flow observations which the logarithmic transformation discards and which constitute a quarter of the sample and the whole of the object of the following chapter. The second is that it does not require the condition on the expectation of the log error which the heteroskedasticity test rejects and which the estimated variance function gives no reason to expect to hold. The third, deferred because it cannot be settled within the estimation sample, is that the Poisson estimator predicts destinations withheld from estimation more accurately in levels.

The specification was assembled by adding regressors one at a time under a functional-form test rather than by searching for the best fit. The procedure required the inclusion of per capita alcohol consumption, without which the core gravity model is misspecified, and it determined the form in which domestic wine production enters, rejecting the logarithmic transformation that fitted better but implied implausible substitution among small producers. The resulting model attributes Prosecco exports to income per head rather than demographic mass, to proximity, to the propensity to consume alcohol, to tariff protection and to the presence of a competing domestic industry.

Two coefficients are less secure than the others, and the chapter has been explicit about both. The tariff elasticity of -3.461 , though close to the median reported in the gravity literature, rests on the twelve destinations whose applied rate varies over time and falls by more than half

when the United States and the United Kingdom are excluded; Chapter 4 nonetheless retains the variable and examines its consequences by simulation rather than by omission. The European Union indicator is negative in the cross-section, but the sign does not survive restriction of the sample to European destinations, the coefficient is what remains after the tariff and agreement channels have already been absorbed, and the composition of the comparison group is sufficient to account for it. It is retained as an uninterpreted control and is not read as an effect of membership.

The comparison with Echevarria Garay et al. (2025) shows that the dataset behaves as the published record would lead one to expect within the range of markets that study covers and that the divergence between the two works is one of question rather than of result. The analysis of the 2025 United States tariff, conducted on monthly data with the other leading markets as controls, finds clear evidence of anticipatory stockpiling after the presidential election and no detectable contraction in the eight months of full treatment, while establishing that the design has power to exclude large effects but not small ones. Chapter 4 uses the model to compare predicted with realised exports across the full set of destinations, and takes up the question of which markets absorb less Prosecco than their fundamentals would suggest.

Chapter 4

Unexploited export potential

4.1 Measuring potential out of sample

The model estimated in Chapter 3 explains Italian Prosecco exports as a function of the economic size of each destination, its distance from Italy, its propensity to consume alcohol, the tariff it applies and the domestic wine industry it possesses. For every destination and year it therefore produces a predicted value: the level of exports that a market with those characteristics would be expected to absorb, given the average relationship observed across all destinations. Comparing that prediction with what was actually exported yields a residual; the interpretation of that residual is the subject of this chapter. The predictions used throughout are those of the preferred specification of Section 3.4, estimated on the panel of 211 destinations described in Chapter 2, and they incorporate the tariff each market applies as documented in Section 2.8.

The interpretation is not automatic. A market importing less than predicted may be doing so because something is holding it back — an absent distributor, an unfamiliar consumer, a regulatory obstacle not captured by the tariff variable — or simply because the model is imperfect and the prediction is wrong. The two possibilities are observationally identical in a single cross-section and the difference between an analysis of export potential and an arbitrary ranking of residuals lies entirely in how they are distinguished.

The approach adopted here rests on a distinction between two uses of the same model. To estimate how large a gap must be before it can be read as evidence rather than noise, the model must be asked to predict markets it has not seen; that is the purpose of the out-of-sample design set out below, and it yields a threshold. To rank the markets themselves, the model estimated on the full sample is used, because that is the most precise estimate of the expected level available. The threshold determines which gaps are legible; the full-sample estimates determine what they are.

Two further points of definition are needed. The analysis is conducted on export value in euros rather than on quantity, since the object of interest is commercial opportunity; results on quantity are similar throughout, as Chapter 3 showed. And the gaps are expressed as annual averages over the eight years of the panel rather than for a single year, which avoids attributing to a market a shortfall arising from one anomalous season.

The difficulty with using an estimated model to identify under-served markets is that the model

has seen those markets. Its coefficients were chosen partly to fit them, so the predicted value for any destination in the estimation sample incorporates that destination's own contribution. A residual computed in this way understates how far the model would miss a market whose behaviour it had not been able to accommodate.

The remedy is to withhold destinations from estimation and predict them afterwards. The 211 destinations are divided into five groups. The model is estimated on four of them and used to predict the fifth and the operation is repeated until each group has served once as the prediction set. Every destination thereby receives a prediction from a model that has never observed it, including the thirty-nine markets that have never imported. The whole procedure is repeated ten times with different assignments, since a single division would make the result depend on which country happened to fall where, and the reported figures are averages across the fifty resulting estimations.

Two features of the division deserve comment. The groups are formed among destinations rather than among observations: a random split at the level of the country-year cell would place Germany in 2019 in the estimation set and Germany in 2020 in the prediction set and the model would already have learned what it was being asked to predict. And the division is stratified by market size, the destinations being ordered by cumulative export value and sampled from within each fifth, because an unstratified split can produce a group composed entirely of minor markets, whose results are then uninterpretable.

The choice of error measure is equally consequential. Mean squared error is unusable here: with flows ranging from zero to 474 million euros it would be determined almost entirely by the United States and the United Kingdom and would say nothing about the markets the analysis is concerned with. Percentage errors are undefined at precisely the zero-flow observations that motivate the exercise. The measures used are therefore: the Poisson deviance, which accommodates zeros by construction and is the loss function the estimator itself minimises; the rank correlation between predicted and observed values, which asks whether the ordering of markets is right independently of the levels; the ratio of predicted to observed aggregate, which measures calibration; and the median absolute deviation between predicted and observed values in logarithms, which expresses the typical error as a multiplicative factor.

Table 4.1 reports the outcome, with the log-linear specification of Chapter 3 estimated alongside as a point of comparison.

Table 4.1: Out-of-sample predictive performance

	PPML	OLS on logs
Poisson deviance per observation (€ mn)	9.08	10.09
Rank correlation with observed	0.766	0.789
Calibration (predicted / observed)	1.21	1.39
Typical multiplicative error	2.71×	2.74×

Averages over ten repetitions of five held-out groups of destinations, stratified by market size. The log-linear model is estimated on positive flows within each estimation set and retransformed with the Duan correction. Source: own elaboration.

The Poisson estimator is clearly superior on the two measures concerning levels. Its out-of-sample deviance is a tenth lower, and it overstates aggregate exports by 21% against 39% for the log-linear model — a difference that matters directly, since a systematic tendency to over-predict inflates every gap and would generate apparent potential everywhere. On the ordering of markets the two are close, with a slight advantage to the log-linear specification at 0.789 against 0.766. The advantage is small and it does not overturn the choice made in Chapter 3, since the estimation of potential requires levels and not merely an ordering, but it is reported here rather than omitted.

The final row is the figure on which the rest of the chapter depends. For a destination it has not seen, the model is typically wrong by a factor of 2.71 in one direction or the other. This is a large error, and it should be. The model contains eleven regressors and is asked to predict the Prosecco imports of a country about which it knows only its income, its population, its distance from Italy, its drinking habits, its tariff and its wine industry; that it comes within a factor of three is perhaps the more notable of the two.

The consequence for what follows is a screening rule. A gap between predicted and realised exports is retained for examination only when it exceeds that factor: a market importing half of what the model predicts is entirely unremarkable, a market importing a tenth is not. The nature of the criterion should be stated precisely. The figure of 2.71 is the average, across the fifty rotations, of the median multiplicative error observed when the model predicts destinations withheld from estimation. It is not a confidence interval and it is not a bound established by inference, so a gap exceeding it is not thereby shown to be statistically distinguishable from prediction error. It is an empirical screening criterion calibrated on the model's own out-of-sample performance, and its merit is that it was measured rather than chosen: the alternative would have been a cut-off selected by the analyst, open to the objection that a different one would have produced a different list. A stricter criterion could be defended. The median error is itself stable across rotations, ranging from 2.61 to 3.62 between the lowest and the highest of the fifty, and the upper end of that range is the natural candidate for a conservative variant. Applying 3.62 instead of 2.71 retains six of the eleven markets and 108 of the 142 million euros: China, Ireland, South Korea, Brazil, Slovenia and Israel survive, while Japan, Denmark, Luxembourg, Iceland and Albania fall below. The markets at the head of the ranking are therefore insensitive to the choice; those at its foot are not, and the gaps reported for them should be read with that in mind.

The same design can be turned on the specification itself. Chapter 3 assembled it by adding regressors under a functional-form test; a criterion making no reference to that test is whether each successive specification predicts better the destinations it has not seen.

Table 4.2: The four specifications compared out of sample

Specification	Deviance € mn	Rank correlation	Calibration	Typical error
(1) Core gravity	12.41	0.715	1.35	3.30×
(2) + alcohol consumption	9.24	0.765	1.23	2.74×
(3) + tariff	8.64	0.767	1.20	2.84×
(4) + domestic production	9.08	0.766	1.21	2.71×

Averages over ten repetitions of five held-out groups of destinations, stratified by market size. All four estimated by PPML with year fixed effects. Source: own elaboration.

One conclusion is clear and one is not. The core gravity specification is markedly the worst on every measure: it misses an unseen market by a factor of 3.30 where the others miss by roughly 2.7, and it overstates the aggregate by 35%. Adding per capita alcohol consumption improves all four measures at once. The functional-form test of Chapter 3 and the out-of-sample criterion therefore agree on the point that matters most, which is that a gravity model of Prosecco exports ignoring how much alcohol a destination drinks is not merely misspecified but appreciably less accurate.

Among the remaining three specifications the criterion does not discriminate. The differences lie in the second decimal place and they are not stable: repeating the exercise with different random assignments of destinations to groups reverses the ordering between the last two. This is informative in itself, since it indicates that the out-of-sample criterion cannot settle the remaining choices. Those rest instead on the functional-form test and on what each variable is required to do. Domestic production is retained because Section 4.4 shows that its omission introduces wine-producing countries into the ranking that do not belong there; the tariff is retained because the counterfactual of Section 4.4 cannot be constructed without it.

These results also bear on the choice of estimator left open in Chapter 3. The comparison conducted there rested on properties holding within the estimation sample, where each estimator reproduces by construction the aggregate of the quantity it fits, and such a comparison cannot discriminate between them. The comparison conducted here is of a different kind, since no first-order condition governs the behaviour of either model on destinations it has not seen. The evidence it provides is favourable to the Poisson estimator but should not be overstated, and the reason lies in the criteria themselves.

The deviance is the loss function the Poisson estimator minimises, so its advantage on that measure is close to being assumed. Calibration is related to the same first-order condition, and although nothing guarantees it out of sample, the connection is not remote. The two neutral measures are the rank correlation and the typical multiplicative error, and on both the estimators are nearly tied, at 0.766 against 0.789 and 2.71 against 2.74.

The honest reading is that out-of-sample performance does not favour the log-linear model and favours the Poisson estimator on the measures concerning levels, without settling the matter on its own. What settles it are the two arguments of Chapter 3 that do not depend on any comparison of fit: the log-linear model cannot use a quarter of the sample and its consistency requires a condition on the expectation of the log error that the heteroskedasticity test rejects. The out-of-sample evidence is consistent with that choice and adds the practical assurance that the model used in the remainder of this chapter is calibrated to within 21% of the aggregate on markets it has never seen.

4.2 The markets never served

The panel constructed in Chapter 2 covers the universe of World Bank economies rather than only those Italy has supplied, and thirty-nine destinations record no Prosecco imports in any year of the period. These are the natural first candidates for unexploited potential and the rotation design produces a prediction for each of them from a model that has never observed

them. Their combined predicted absorption is 5.16 million euros a year, against average annual exports of 1384 million. The entire set of markets Prosecco has never reached is therefore worth under four tenths of one per cent of current exports.

This is the first result of the chapter, and it is a negative one. It becomes more informative when the thirty-nine are examined individually, since most of that small total does not represent commercial opportunity at all. They divide into three groups, and the division is not a matter of degree but of kind.

Table 4.3: The thirty-nine never-served markets by category

Category	Markets	Predicted € mn/year	Share of the 39	Share of exports
Recorded under another jurisdiction	6	2.59	50.1%	0.19%
Sale prohibited or absent	13	2.36	45.8%	0.17%
Genuinely unexplored	20	0.21	4.1%	0.015%
Total	39	5.16	100%	0.37%

Predicted values are out-of-sample averages from the rotation design. Classification based on national legislation on alcohol and on the customs status of each territory. Source: own elaboration.

The first group comprises six territories whose trade is recorded under another customs jurisdiction: Monaco within the French customs territory, San Marino in customs union with Italy, the Channel Islands and the Isle of Man within the United Kingdom, Puerto Rico within the United States and Saint-Martin within France. Prosecco reaches these places; it is simply attributed elsewhere in the statistics. Their zeros are a recording convention rather than an absence of demand, and the 2.59 million euros the model predicts for them is in all likelihood already being exported and counted under a different heading. Monaco alone accounts for 1.25 million of it. A corroborating detail is that these are also, with few exceptions, the territories for which the World Health Organization reports no separate alcohol consumption figure, the same statistical convention operating in another dataset.

The second group comprises thirteen markets in which the sale of alcohol is prohibited, restricted to non-residents, or culturally absent without formal prohibition: Saudi Arabia, Iran, Yemen, Afghanistan and Mauritania under total prohibition; Pakistan, Sudan, Brunei and the Comoros with exemptions for non-Muslims or foreign residents; Algeria, Syria, Niger and the West Bank, where sale is not forbidden but consumption is negligible. The model predicts 1.35 million euros a year for Saudi Arabia because it observes an economy of over one thousand billion dollars and does not observe that the product cannot legally be sold there. This is a limitation of the specification rather than a finding about the market, and it is reported as such.

What remains is twenty destinations that are genuinely open and genuinely unserved, and they are worth 0.21 million euros a year in total — fifteen thousandths of one per cent of current exports. The largest are the Faroe Islands and Greenland at some sixty and fifty thousand euros, followed by Papua New Guinea, Eswatini and Tajikistan. These are small, low-income or very distant markets, and the model's assessment of them is that a bottle of Prosecco would find few buyers. Nothing in the data suggests otherwise.

The conclusion is therefore that the markets Prosecco has never reached contain no meaningful commercial potential and that the greater part of what appears as potential among them is

explained by statistical convention and by legal prohibition. The question was a legitimate one, and the answer redirects the analysis rather than closing it.

4.3 Under-served active markets

If potential does not lie among the destinations Prosecco has never reached, it may lie among those it has reached but in quantities below what their characteristics would suggest. Applying the same comparison to markets with positive flows, and retaining those whose predicted value exceeds the realised value by more than the out-of-sample error factor of 2.71, yields eleven destinations. The threshold on cumulative exports of one million euros over the period excludes markets too small for the comparison to be meaningful. It removes 81 of the 172 destinations with positive flows, but those 81 account for two tenths of one per cent of export value; the largest among them imports 124,000 euros a year and the median 21,000, quantities at which the ratio of predicted to realised value is governed by rounding and by occasional shipments rather than by market conditions. The choice is not delicate: raising the cut-off to two million or lowering it to half a million leaves 99.7% and 99.9% of export value inside the comparison respectively.

Table 4.4: Active markets importing less than predicted

Market	Actual € mn/year	Predicted € mn/year	Ratio	Gap € mn/year
China	2.89	35.94	12.4×	33.05
Ireland	4.82	33.88	7.0×	29.07
South Korea	1.35	20.33	15.1×	18.98
Japan	6.22	22.29	3.6×	16.07
Denmark	5.37	17.91	3.3×	12.54
Brazil	0.67	12.87	19.3×	12.21
Slovenia	0.99	11.59	11.7×	10.59
Israel	0.85	5.19	6.1×	4.35
Luxembourg	1.09	3.33	3.1×	2.25
Iceland	0.89	2.65	3.0×	1.76
Albania	0.30	0.94	3.1×	0.64
Total	25.4	167.0		142

Markets with annual exports above one million euros whose predicted value exceeds the realised value by more than the out-of-sample error factor of 2.71. Predictions from the preferred PPML specification of Chapter 3 estimated on the full sample. Figures are annual averages over 2018–2025. Source: own elaboration.

The eleven markets together account for a gap of 142 million euros a year, equivalent to 10.2% of current Italian Prosecco exports. Against the 5.16 million of the never-served markets the difference is of a different order: unexploited potential is concentrated among destinations that already buy the product, by a factor of some twenty-seven to one. This is the principal finding of the thesis and it carries a straightforward implication for commercial strategy: the question is not where Prosecco has yet to arrive, but where it has arrived and failed to establish itself.

This figure describes the potential available under present trade policy, since the predictions incorporate the tariff each market actually applies. Section 4.4 reports the corresponding calcu-

lation under a counterfactual in which tariffs are set to zero, and shows that the two questions have different answers and identify partly different markets.

The composition of the list rewards examination. The three largest gaps are in Asia and Ireland. China and South Korea are markets of considerable size and rising income where Prosecco imports remain, respectively, 2.9 and 1.4 million euros a year — figures comparable to those of Latvia, a country of 1.8 million inhabitants. Japan, a mature market for imported wine, imports 6.2 million against a predicted 22.3. Brazil, at 0.7 million against a predicted 12.9, shows the largest proportional gap of the group.

Ireland and Denmark are of a different kind. Both are wealthy, proximate, high-consuming European markets without domestic production, and the model predicts substantial absorption on those grounds alone. Ireland imports 4.8 million euros a year against a predicted 33.9. That a market so favourably placed should fall so far short is the least easily explained result in the table, and Section 4.5 returns to it.

Slovenia is the one entry whose presence should be treated with caution. It is a contiguous, high-income market, but it produces 24.9 litres of wine per inhabitant, and although the specification controls for domestic production the control uses total wine output rather than sparkling wine specifically. Slovenian production is oriented towards still wines, so the control may understate the extent to which domestic supply competes with the imported product; alternatively the gap may be genuine. The data available do not resolve the question.

The ranking rests on predictions from the model fitted to the whole sample, while the threshold that screens it was calibrated out of sample. The two can be brought together by applying the same rule to the rotation predictions themselves, so that each market is judged against a model that never observed it. Ten of the eleven markets survive, the rank correlation between the two sets of gaps is 0.988, and the two markets on which the lists disagree, Iceland and Belarus, sit within four hundredths of the threshold. The identification of the under-served markets therefore does not depend on the model having seen them.

Two differences are worth recording. The gaps measured out of sample are larger throughout, 205 against 142 million euros once Germany is set aside, which follows from the calibration reported in Table 4.1: rotation predictions overstate the aggregate by roughly a fifth. This is the reason the ranking is built on the full-sample model rather than on the rotation, since what is wanted is the best available estimate of the expected level, not a deliberately weakened one.

Germany is the second difference and a more substantial one. Withheld from estimation it is predicted at 311 million euros a year against 72 realised, a gap of 239 million that alone accounts for most of the distance between the two totals. Germany is among the largest destinations, and the Poisson estimator weights observations by size; when the market is in the estimation sample the coefficients accommodate it, and when it is withheld they no longer do. The figure measures the influence of a single large market on the estimated coefficients rather than an unexploited potential of that magnitude, and it is reported here for that reason rather than added to the ranking.

A second check addresses the European Union indicator, which carries a large negative coefficient and might be standing in for geographical heterogeneity that the specification does not otherwise capture. Replacing it with macro-region fixed effects asks each market to be judged against the others in its own region rather than against all destinations. Eight of the eleven markets survive, among them China, South Korea, Ireland, Slovenia, Brazil and Israel, and the aggregate gap of 100 million euros a year remains twenty times the potential of the markets

never served. Japan, Denmark and Luxembourg drop out; Norway, Belarus, Serbia, Portugal and Montenegro enter; the rank correlation between the two sets of gaps is 0.857.

The list is therefore less stable under this check than under the previous one, and the differences are intelligible as a change in the comparison group rather than as instability in the method. Japan ceases to appear because it is not under-served relative to other Asian destinations, and Norway appears because it is under-served relative to other western European ones. The specification with regional effects is not adopted as the principal one for a reason visible in the estimates themselves: the coefficient on distance falls from -0.928 to -0.140 and loses significance, since regional effects absorb most of the geographical variation on which a gravity equation depends. The exercise is informative about the sensitivity of the ranking to the reference group; it does not provide an alternative estimate of the expected level.

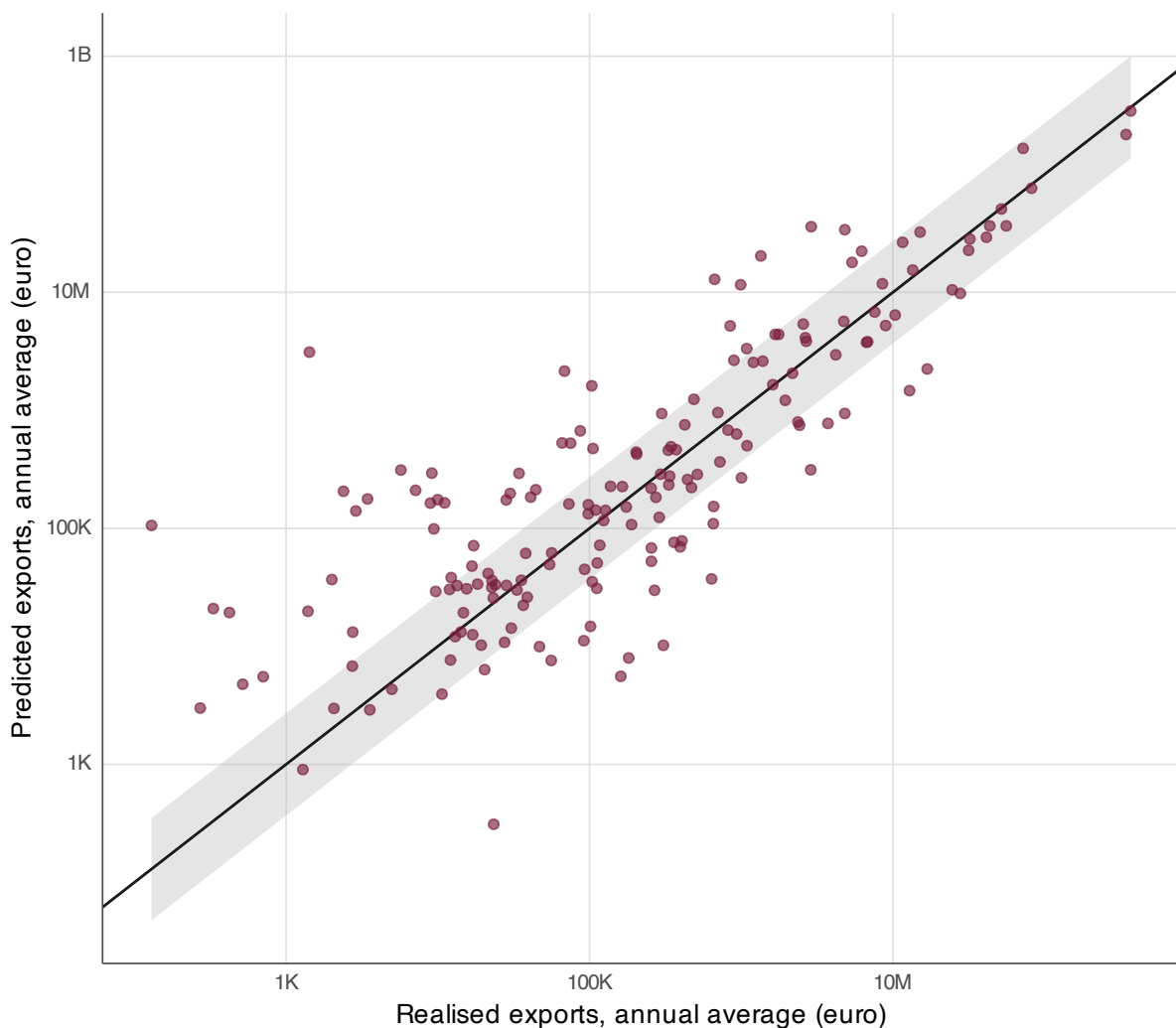


Figure 4.1: Predicted against realised Prosecco exports, annual averages 2018–2025, 172 destinations with positive flows.

Predictions from the preferred PPML specification estimated on the full sample. The solid line marks equality between prediction and observation; the shaded band spans the median out-of-sample error factor of 2.71 established in Section 4.1, within which gaps are not retained for examination. The thirty-nine markets that never import are excluded, since they cannot be placed on a logarithmic scale, and are treated separately in Table 4.3.

Source: own elaboration.

4.4 The contribution of tariffs

Chapter 3 identified two coefficients as insecure: the tariff, whose identification rests on twelve destinations, and the European Union indicator, whose cross-sectional sign does not survive restriction of the sample to European markets. Both enter the predictions on which Table 4.4 rests. The two require different treatment, however, because the questions they raise are not the same.

The tariff raises a question of design rather than of sensitivity. It is tempting to remove the variable from the specification on the ground that the resulting predictions would describe what each market would absorb under equal conditions. That reasoning does not hold. Omitting a variable whose coefficient is significant does not equalise conditions across markets: the tariff differences remain in the data that generated the observed flows and the model simply ceases to observe them, so that their effect is absorbed by the other coefficients or by the residual. The exercise would introduce omitted variable bias rather than construct a counterfactual.

The question is nonetheless legitimate and it can be answered properly by retaining the variable and varying its value. Two sets of predictions are therefore reported. The first uses the tariff each market actually applies and measures the potential available under present trade policy. The second sets the applied tariff to zero for every destination while leaving all other characteristics unchanged and measures the potential that would exist if trade policy ceased to differentiate between markets. In aggregate the second exceeds the first by 19%, predicted absorption rising from 1384 to 1645 million euros a year. The difference between the two, market by market, is the part of the gap associated with protection.

Table 4.5: Potential under current tariffs and under a zero-tariff counterfactual

Market	Applied tariff	Actual € mn/yr	Gap at current tariffs	Gap at zero tariffs	Associated with the tariff
Switzerland	40.1%	51.90	−1.12	112.11	113.23
China	14.0%	2.89	33.05	53.67	20.62
Ireland	0.0%	4.82	29.07	29.07	0.00
South Korea	1.3%	1.35	18.98	19.91	0.93
Japan	3.6%	6.22	16.07	18.52	2.45
Denmark	0.0%	5.37	12.54	12.54	0.00
Brazil	0.0%	0.67	12.21	12.21	0.00
Slovenia	0.0%	0.99	10.59	10.59	0.00
Israel	12.0%	0.85	4.35	6.84	2.49
Turkey	50.0%	1.61	0.03	5.09	5.06
Belarus	12.5%	1.76	2.64	4.86	2.22
India	150.0%	0.66	−0.50	3.00	3.50
Luxembourg	0.0%	1.09	2.25	2.25	0.00
Iceland	0.0%	0.89	1.76	1.76	0.00
Albania	0.0%	0.30	0.64	0.64	0.00
Total			142	293	150

Predictions from the preferred PPML specification estimated on the full sample. The counterfactual sets $\ln(1+\tau)$ to zero for every destination, holding all other regressors at their observed values. Markets qualify if their annual

exports exceed one million euros and the predicted value exceeds the realised value by more than the out-of-sample factor of 2.71 under either scenario. Annual averages over 2018–2025. Source: own elaboration.

Under current tariffs the gap totals 142 million euros a year across eleven markets, as reported in Table 4.4. Under the zero-tariff counterfactual it rises to 293 million across fifteen, and the four additional entrants — Switzerland, Turkey, Belarus and India — are precisely the markets whose present protection is high enough to suppress the flows the model would otherwise predict.

Switzerland is the case worth setting out, and it does not appear in the analysis at current tariffs at all. With a duty of 40.1% it imports 51.9 million euros a year, almost exactly what the model predicts for it; there is no anomaly to explain. Set the tariff to zero in the model and predicted absorption rises to 164.0 million. The figure is a scenario rather than a forecast: it states what the estimated relationship implies for a market with Switzerland's income, population, distance and drinking habits facing no duty, holding everything else at its observed value. It is not a claim about what Switzerland would import were the duty removed, which would depend on adjustments the model does not represent, among them the distribution capacity of a market that has had no reason to develop one at the current level of protection. Read as a scenario it says something definite: the Swiss market is not under-served relative to its circumstances, it is constrained by them, and the binding constraint is a tariff. That is a different finding from the ones in Table 4.4, and one whose implications differ, since trade policy is subject to negotiation in a way that consumer preferences are not.

Two cautions apply to the counterfactual. The first is that it extrapolates. India applies a duty of 150%, and removing it implies multiplying predicted imports by a factor of nearly twenty-four; the tariff coefficient was estimated over a range in which values that high are represented by a single destination, and the resulting figure should be treated as an illustration of the mechanism rather than as a forecast. The second is that the simulation holds everything else constant, including the distribution structures and consumer familiarity that a market with high protection is unlikely to have developed. The counterfactual measures the tariff's contribution to the gap within the model, not the sales that would follow its removal.

The remaining two variables raise ordinary questions of sensitivity and Table 4.6 reports them.

Table 4.6: The ranking under alternative specifications

Specification	Markets	Total gap € mn/year	In common with preferred	Leading market
Preferred	11	142	—	China
Without the EU indicator	10	320	9 of 11	Germany
Without domestic production	16	161	11 of 11	Spain

Each row re-estimates the preferred specification omitting one variable and recomputes the list of under-served markets on the same criteria. Source: own elaboration.

The European Union indicator matters a great deal. Omitting it more than doubles the estimated total gap, from 142 to 320 million euros a year, and places Germany at the head of the list. The

reason is the one given in Section 3.4: the indicator absorbs a regional effect that the gravity variables do not capture, and without it the model predicts for European destinations a level of imports that European destinations systematically do not reach. Germany's apparent shortfall in that specification is not evidence about the German market but evidence that the model, deprived of the regional control, over-predicts European markets as a class. The variable is retained for this reason and the figures reported in Table 4.4 are the conservative ones.

Domestic production affects the composition of the list rather than its size. Without it, sixteen markets qualify and Spain heads the ranking, followed by Portugal and Chile. All three are substantial wine producers, and their apparent shortfall reflects a domestic industry supplying a close substitute rather than an untapped market for Prosecco. The eleven markets of Table 4.4 all survive the omission, which indicates that the control removes false positives without introducing new ones.

4.5 Summary and qualifications

The gaps in Table 4.4 measure the difference between what a market imports and what a model predicts it would import given its characteristics. They do not identify the cause of that difference, and four possibilities remain open in each case.

The first is that the model omits a relevant characteristic. Chapter 3 documented the variables it contains; consumer preferences over sparkling wine styles, the structure of alcohol distribution, excise duties and the presence of competing imported products are not among them. A market may import less than predicted because it prefers a different product and the model has no way of registering the difference.

The second is commercial rather than economic. Establishing a position in a distant market requires a distributor, a listing in the retail chains and a period of investment in brand recognition. A gap may therefore record the absence of a commercial relationship rather than the absence of demand, which is precisely the kind of gap a consortium can act upon — but the data cannot distinguish it from the first case.

The third is that the gap is a prediction error. The out-of-sample threshold used here excludes gaps below a factor of 2.71, but it is a median: individual markets can and do lie further from their predicted values without any economic explanation being required. The markets at the bottom of Table 4.4, whose ratios lie close to three, are the more exposed to this reading.

The fourth applies particularly to Ireland and Luxembourg. Both are small, wealthy economies with disproportionately large trade in transit, and imports recorded at their borders may not be consumed within them. Latvia, which does not appear in the table but records the highest per capita Prosecco imports in Europe at 8.96 euros per inhabitant, is a clearer instance of the same phenomenon: for a country of 1.8 million people, that figure is more plausibly explained by re-export than by domestic consumption.

The gravity model treats recorded imports as final demand and where that assumption fails the predictions inherit the error. None of these considerations invalidates the ranking, but together they determine how it should be read: as a set of markets warranting investigation, ordered by the size of the anomaly, rather than as a measurement of achievable sales. The figure of 142 million euros a year is the sum of the anomalies, not a forecast of revenue that a commercial effort would capture.

This chapter has used the model of Chapter 3 to compare predicted with realised Prosecco exports across the full set of destinations. The comparison was calibrated by an out-of-sample design in which destinations are withheld from estimation in rotation, which established that the model misses an unseen market by a median factor of 2.71 and thereby supplied the screening criterion applied to the gaps.

The thirty-nine markets that have never imported Prosecco are worth 5.16 million euros a year in predicted absorption, under four tenths of one per cent of current exports, and the greater part of that figure is accounted for by six territories whose trade is recorded under another jurisdiction and thirteen markets in which the sale of alcohol is prohibited or absent. Genuinely unexplored and genuinely open markets are worth 0.21 million euros a year in total. There is no unexploited potential in the destinations Prosecco has not reached.

There is a considerable amount among those it has. Eleven active markets import less than a third of what their characteristics predict, for a combined gap of 142 million euros a year, equivalent to a tenth of current exports. China, Ireland, South Korea, Japan, Denmark and Brazil head the list. The ranking depends materially on the retention of the European Union indicator, without which the model over-predicts European markets as a class.

Retaining the tariff in the specification and setting it to zero in simulation answers a further question: how much of the shortfall is associated with trade policy rather than with commercial factors. Under that counterfactual the gap rises to 293 million euros a year across fifteen markets, and four destinations enter the list that do not appear under current tariffs. Switzerland is the clearest case: it imports almost exactly what the model predicts given a duty of 40.1%, and the zero-tariff scenario implies three times that level for a market with its characteristics and no duty.

The distinction matters, since a shortfall caused by protection and a shortfall caused by the absence of a distribution network call for entirely different responses. The gaps identify markets warranting investigation rather than sales that could be captured, and the reasons for them lie beyond what a gravity model can establish. The concluding chapter considers what follows for the strategy of the sector and what further evidence would be required to distinguish between the explanations that remain open.

Conclusions

The question and the answer

This thesis set out to establish where Italian Prosecco exports fall short of what the characteristics of foreign markets would lead one to expect, and to quantify that shortfall. The question was prospective: it concerned not the destinations that already absorb the product but those that, on the evidence of their income, their proximity, their drinking habits and the barriers they impose, might absorb more of it than they do. The answer is that the shortfall exists, that it amounts to approximately 142 million euros a year and that it lies in a different place from where the sector's own strategic vocabulary would suggest looking.

Diversification is usually discussed in terms of new markets and the appellation's export strategy has long been framed around reaching destinations where Prosecco is not yet present. The evidence assembled here does not support that framing. The thirty-nine markets Prosecco has never reached are worth, in aggregate, 5.16 million euros a year of predicted absorption, under four tenths of one per cent of current exports; and of that small figure, half is accounted for by six territories whose trade is recorded under another customs jurisdiction and where Prosecco already arrives, and a further 46% by thirteen markets in which the sale of alcohol is prohibited or culturally absent. What remains genuinely unexplored and genuinely open is worth 0.21 million euros a year: an amount that rounds to zero at any scale relevant to commercial decision-making.

The potential lies instead among destinations that already import Prosecco but in quantities far below what their characteristics predict. Eleven such markets were identified, applying a screening criterion measured rather than chosen: the model, asked to predict destinations withheld from estimation, misses them by a median factor of 2.71, and only gaps exceeding that factor were retained. China, Ireland, South Korea, Japan, Denmark and Brazil head the list. Their combined shortfall is equivalent to a tenth of current exports and it exceeds the potential of the never-served markets by a factor of twenty-seven. The pattern suggests a reframing of the question: not where Prosecco has yet to arrive, but where it has arrived and has not established itself. The evidence supports the reframing; it does not by itself explain why those markets absorb less than their characteristics would suggest.

Implications for the sector

Three consequences follow for anyone concerned with the commercial strategy of the appellation and they should be read as indications of where to look rather than as recommendations, for reasons set out below. The first concerns the allocation of effort. If the markets Prosecco

has never reached are worth two tenths of one per cent of current exports at best, the resources devoted to opening them are difficult to justify against the alternative of deepening penetration where a commercial presence already exists. The two largest gaps identified here, China and Ireland, are together worth 62 million euros a year, more than ten times the entire potential of the unserved set.

The second concerns the distinction between commercial and policy obstacles, which the analysis of tariffs makes it possible to draw. A market may import less than predicted because no distributor has established a position there, or because a duty makes the product uncompetitive; the two call for entirely different responses, and the counterfactual reported in Section 4.4 separates them. Switzerland is the instructive case. With a duty of 40.1% it imports almost exactly what the model predicts and by the criterion of Chapter 4 it is not under-served at all; set the duty to zero in the model and predicted absorption rises from 51.9 to 164.0 million euros a year. The second figure is a scenario, not a forecast of what removing the duty would deliver: it describes what the estimated relationship implies for a market with Switzerland's characteristics facing no tariff, and says nothing about the distribution capacity such a market would need to build. What it establishes is a distinction: the Swiss market is not neglected but constrained, and what constrains it is negotiable in a way that consumer preferences are not. Turkey, Belarus and India fall into the same category on a smaller scale.

The third concerns exposure. Exports remain concentrated, though markedly less so than at the beginning of the period: the Herfindahl index of destination shares fell from 1953 in 2018 to 1320 in 2025, in every single year, and the combined share of the two leading markets from 59.1% to 47.7%. That the largest of them subsequently imposed a tariff and that the sector's exposure to that decision was proportionate to its dependence on a single destination, is an argument for the diversification already under way rather than evidence against it. The analysis of the 2025 measure found no detectable contraction in flows over the eight months of full application, but the design has power to exclude only large effects, and the absence of a visible response should not be mistaken for the absence of risk.

What the analysis cannot establish

A gravity model explains trade by the characteristics it observes and the interpretation of what it fails to explain is constrained by what it does not observe. Four limitations bear directly on how the results of this thesis should be read, and they are restated here because the figures reported above are of a kind that invites over-interpretation.

The gaps identify anomalies, not causes. A market importing a tenth of what the model predicts may be doing so because consumers there prefer a different style of sparkling wine, because the distribution chain has no incentive to carry the product, because excise duties or licensing rules raise the final price beyond what the tariff variable captures or because the prediction is simply wrong. The analysis cannot distinguish between these and the figure of 142 million euros a year is the sum of the anomalies rather than a forecast of revenue that commercial effort would capture.

Some of the anomalies may be artefacts of how trade is recorded. Ireland and Luxembourg are small, wealthy economies with substantial trade in transit, and imports crossing their borders need not be consumed within them. Latvia, which does not appear among the under-served markets but records the highest per capita Prosecco imports in Europe at 8.96 euros per inhab-

itant, is a clearer instance of the same phenomenon: for a country of 1.8 million people that figure is more plausibly explained by re-export than by domestic consumption. The gravity model treats recorded imports as final demand and where that assumption fails the predictions inherit the error.

Two coefficients in the model are less secure than the others, and neither supports the interpretation that would be natural to place on it. The tariff elasticity of -3.461 is close to the median reported in the gravity literature, but it is identified almost entirely from comparisons between countries rather than from observed changes in commercial policy, since the applied rate varies over time for twelve destinations only; the counterfactual of Section 4.4 rests on it and inherits its fragility, particularly in the case of India, where removing a duty of 150% requires an extrapolation far beyond the range over which the coefficient was estimated. The European Union indicator is negative in the cross-section, but the sign does not survive restriction of the sample to European destinations, and what remains attached to it after income, distance, alcohol consumption, tariffs and trade agreements have been accounted for is a residual without natural interpretation. It is retained as a control and no conclusion is drawn from it.

Finally, the single-exporter design precludes the multilateral resistance terms of the Anderson and Van Wincoop (2003) formulation, since destination fixed effects would absorb the trade-cost variables of interest. The distance elasticity reported here therefore incorporates whatever correlation exists between bilateral distance from Italy and the general remoteness of each destination from world markets and is not directly comparable with estimates obtained from multi-country specifications.

Directions for further work

Three extensions would strengthen the analysis and each addresses a limitation identified above. The most immediate concerns the substitution control. Domestic wine production enters the model as total output per head, whereas the product that competes with Prosecco is sparkling wine specifically. The distinction matters for at least one market in the ranking: Slovenia produces 24.9 litres of wine per inhabitant, largely still wine, and whether its apparent shortfall survives a properly specified substitution control cannot be settled with the data used here. Country-level series on sparkling wine output would resolve it.

The second concerns the identification of tariff effects. The monthly analysis of the 2025 United States measure was conducted eight months after entry into force, and the design was shown to have power only for effects of considerable size. Repeating it with two or three further years of data, and extending the approach to the other episodes of tariff change recorded in the panel, would allow the cross-sectional elasticity to be checked against a policy-based estimate rather than merely set alongside it.

The third concerns the nature of the gaps themselves. Establishing whether a shortfall reflects an absent distribution network, a mismatch of consumer preferences or a recording artefact requires evidence of a kind that trade statistics do not contain: firm-level export records, distribution agreements, or survey data on consumer awareness in the markets concerned. The eleven destinations identified here would be a natural starting point for such an investigation and the contribution of a gravity analysis is precisely to narrow the field to a number of markets that can be examined individually.

The thesis has tried throughout to distinguish what the evidence supports from what it merely suggests, and to state the limits of each claim alongside the claim. The central finding is robust to every sensitivity examined: the potential for Italian Prosecco lies not in the markets it has yet to reach, but in those it has reached and not yet taken.

References

- Agostino, M. and Trivieri, F. (2014). Geographical indication and wine exports. An empirical investigation considering the major European producers. *Food Policy*, 46:22–36.
- Anderson, J. E. and Van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review*, 93(1):170–192.
- Anderson, K. and Pinilla, V. (2022). Wine’s belated globalization, 1845–2025. *Applied Economic Perspectives and Policy*, 44(2):742–765.
- Barisan, L., Pomarici, E., Boatto, V., and Galletto, L. (2016). Roles of the vine-growing co-operation in the development of a successful wine: The DOCG Prosecco wine co-operatives’ model. In *International Co-Operative Alliance Research Conference*, Almería, Spain. Conference presentation.
- Breusch, T. S. and Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47(5):1287–1294.
- Castillo, J. S., Villanueva, E. C., and García-Cortijo, M. C. (2016). The International Wine Trade and Its New Export Dynamics (1988–2012): A Gravity Model Approach. *Agribusiness*, 32(4):466–481.
- Chen, J. and Roth, J. (2024). Logs with Zeros? Some Problems and Solutions. *The Quarterly Journal of Economics*, 139(2):891–936.
- Dal Bianco, A., Boatto, V. L., Caracciolo, F., and Santeramo, F. G. (2016). Tariffs and non-tariff frictions in the world wine trade. *European Review of Agricultural Economics*, 43(1):31–57.
- del Rey, R. and Mueller Loose, S. (2025). State of the international wine markets in 2024: Stability in global wine trade amid declining consumption and production. *Wine Economics and Policy*, 14(2):119–132.
- Duan, N. (1983). Smearing Estimate: A Nonparametric Retransformation Method. *Journal of the American Statistical Association*, 78(383):605–610.
- Echevarria Garay, J., Stiletto, A., and Trestini, S. (2025). Prosecco Under Pressure: Modelling Export Flows and U.S. Tariff Scenarios Using a Gravity Model Approach. Preprint, Preprints.org. <https://doi.org/10.20944/preprints202507.1963.v1>.
- Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Theory. *Econometrica*, 52(3):681–700.
- Head, K. and Mayer, T. (2014). Gravity Equations: Workhorse, Toolkit, and Cookbook. In *Handbook of International Economics*, volume 4, pages 131–195. Elsevier.

- Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, 70(5):950–959.
- Macedo, A., Rebelo, J., and Gouveia, S. (2023). The impact of COVID-19 government policy on the international wine trade. *Agricultural and Food Economics*, 11(1):44.
- Mariani, A., Pomarici, E., and Boatto, V. (2012). The international wine trade: Recent trends and critical issues. *Wine Economics and Policy*, 1(1):24–40.
- Mayer, T. and Zignago, S. (2011). Notes on CEPII’s distances measures: The GeoDist database. Working Paper 2011-25, CEPII.
- Meloni, G. and Swinnen, J. (2018). Trade and terroir. The political economy of the world’s first geographical indications. *Food Policy*, 81:1–20.
- OIV (2025). State of the world vine and wine sector in 2024. Technical report, International Organisation of Vine and Wine, Dijon.
- OIV (2026). State of the world wine sector in 2025. Technical report, International Organisation of Vine and Wine, Dijon.
- Onofri, L. and Boatto, V. (2022). “as time goes by”: Econometric estimation of the Prosecco life cycle. Technical Report EERI RP 2022/13, Economics and Econometrics Research Institute, Brussels.
- Onofri, L., Boatto, V., and Dal Bianco, A. (2015). Who likes it “sparkling”? An empirical analysis of Prosecco consumers’ profile. *Agricultural and Food Economics*, 3(1):11.
- Pomarici, E., Barisan, L., Boatto, V., and Galletto, L. (2019). The Prosecco Superiore DOCG industry structure: Current status and evolution over time. In *The Palgrave Handbook of Wine Industry Economics*, pages 421–435. Palgrave Macmillan, Cham.
- Pomarici, E., Corsi, A., Mazzarino, S., and Sardone, R. (2021). The Italian Wine Sector: Evolution, Structure, Competitiveness and Future Challenges of an Enduring Leader. *Italian Economic Journal*, 7(2):259–295.
- Porter, M. E. (1998). Clusters and the new economics of competition. *Harvard Business Review*, 76(6):77–90.
- Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society, Series B*, 31(2):350–371.
- Silva, J. M. C. S. and Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, 88(4):641–658.
- Simonovska, I. and Waugh, M. E. (2014). The elasticity of trade: Estimates and evidence. *Journal of International Economics*, 92(1):34–50.
- Tinbergen, J. (1962). *Shaping the World Economy: Suggestions for an International Economic Policy*. Twentieth Century Fund, New York.

Data sources

- CEPII (2011). GeoDist database. Centre d'Études Prospectives et d'Informations Internationales. http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele.asp
- European Commission (2026). Access2Markets. Directorate-General for Trade. <https://trade.ec.europa.eu/access-to-markets/>
- International Trade Centre (2026). Market Access Map. <https://www.macmap.org/>
- ISTAT (2026). Coeweb: statistiche del commercio estero. Istituto Nazionale di Statistica. Annual and monthly extractions, CN 2204.10.15, August 2026. <https://www.coeweb.istat.it/>
- OIV (2026). Statistical database. International Organisation of Vine and Wine. Extraction of August 2026. <https://www.oiv.int/what-we-do/statistics>
- UN Comtrade (2026). United Nations Commodity Trade Statistics Database. <https://comtradeplus.un.org/>
- World Bank (2026). World Development Indicators. <https://databank.worldbank.org/source/world-development-indicators>
- World Bank (2026). World Integrated Trade Solution (WITS-TRAINS). <https://wits.worldbank.org/>
- World Health Organization (2026). Global Health Observatory: alcohol consumption data. <https://www.who.int/data/gho>

Software

- R Core Team (2024). R: a language and environment for statistical computing, version 4.4.2. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>
- Bergé, L. (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. *CREA Discussion Papers*, 13. [Cited as the canonical reference for the `fixest` package, version 0.13.2.]
- Wickham, H., François, R., Henry, L. and Müller, K. dplyr: a grammar of data manipulation, version 1.2.1. <https://dplyr.tidyverse.org/>