IMT School for Advanced Studies, Lucca
Lucca, Italy

Brexit, Trade Disruptions, and Export Recommendations

PhD Program in System Science

Track in Economics, Networks, and Business Analytics

XXXIV Cycle

By

Víctor Ortiz-Giménez

2024

The dissertation of Víctor Ortiz-Giménez is approved.

PhD Program Coordinator: Ennio Bilancini, IMT School for Advanced Studies Lucca

Advisor: Prof. Massimo Riccaboni, IMT School for Advanced Studies Lucca

Co-Advisor: Prof. Francesco Serti, IMT School for Advanced Studies Lucca

The dissertation of Víctor Ortiz-Giménez has been reviewed by:

Antonio Navas, University of Sheffield

Dolores Añón Higón, Universitat de València

IMT School for Advanced Studies Lucca
2024

To my parents Begoña and Vicente,
to my sister Elba,
and to my beloved Conchi.

Contents

List of Figures

List of Tables

Acknowledgements

This thesis would have never been written without the constant support of my family. Their belief in me fueled my journey through this research, and their curiosity motivated me to pursue a Ph.D. The distance from them proved to be the most challenging aspect of my experience at IMT.

I extend my heartfelt gratitude to my esteemed professors at the University Carlos III of Madrid and the University of Alicante. Their mentorship not only cultivated my economic acumen but also instilled in me a perpetual thirst for knowledge. A special acknowledgment is reserved for Gabriel Pérez-Quirós, whose guidance during my tenure at the Bank of Spain paved the way for my research trajectory.

At IMT, I had the privilege to work under the supervision of Massimo Riccaboni, and Francesco Serti. Their mentorship has been invaluable, shaping both my academic journey and personal growth. Their steadfast support, both professionally and personally, has been instrumental in navigating the challenges encountered along the path of doctoral research. Working with Massimo has been nothing short of enriching, he was encouraging and supportive since my first day at IMT. Working with Francesco has been truly enlightening, seamlessly merging work with friendship and turning obstacles into meaningful achievements. His influence was pivotal in my decision to pursue a doctorate. His constant faith in me and insightful advice had a profound impact on the direction of my research. He will always be my role model for an advisor. I also want to express my gratitude for their work as co-authors, as well as Federico Nutarelli and Marco Dueñas.

The generous support extended by KU Leuven University during my visiting tenure is gratefully acknowledged. Their financial assistance not only facilitated my research but also served as a testament to their commitment to scholarly endeavors. Special appreciation is reserved for Hylke Vandenbussche that invited me to KU Leuven, whose guidance from February 2022 to September 2023 enriched my understanding of International Trade and expanded the horizons of my research. Also at KU Leuven, Johannes Van Biesebroeck provided excellent suggestions that improved this dissertation and, together with Jan De Loecker, helped me better understand the job-market process. To my colleagues-turned-friends, Giorgos Morakis, Na Liu, Philipp Ludwig, and Zheng Honghao, I owe a debt of gratitude for their camaraderie, fostering an environment where intellectual curiosity thrived.

I wish to express heartfelt gratitude to my friends at IMT, who have become akin to family, Alberto Hidalgo, Ángel Valverde, Chiara Bellucci, Erica Ordali, Ilaria Gianstefani, Juanfran Blazquiz, Pablo Marcos-Prieto, Sara Landi, Sedric Zucchiatti, Stefania Miricola, and Surya Soman. Their genuine friendship and unwavering encouragement have infused every moment at IMT with warmth and significance. Their companionship has truly elevated my Ph.D. experience, making it a journey filled with fulfillment and joy. Furthermore, I would like to further extend my appreciation to my friends beyond the IMT community, including Alberto P.B., Alfonso B.B., Antonio T.C, Gema A.F., Jarn V., Jeroen S., Miguel R.B., and Noel V.G.

Finally, I want to thank my dearest one, Conchi Hernández, who has given me strength and love throughout the years of this project.

Abstract

This doctoral dissertation is structured into three chapters, each addressing aspects related to policy evaluation, with a focus on events carrying substantial implications for the economy and international trade. These events include the Brexit referendum, trade disruptions, and the development of export recommendations.

The Economic Cost of a Referendum. The Case of Brexit. This paper estimates how GDP would have behaved in the United Kingdom after the Brexit referendum in the absence of the mentioned poll using the Synthetic Control Method. We contribute to the research on the effects of Brexit by quantifying the macroeconomic cost of this referendum before the actual Brexit has taken place. We find a large and significant negative effect of the Brexit referendum on the GDP of UK. This loss is increasing in time representing, in $2017\,Q4$, 1.71% of the observed GDP of UK.

Assessing the Heterogeneous Impact of Economy-Wide Shocks: A Machine Learning Approach Applied to Colombian Firms. This paper investigates the impact of COVID-19 on Colombian exports, revealing a substantial decline in survival probabilities during 2020. On average, we find that the COVID-19 shock decreased a firm's probability of surviving in the export market by about 20 percentage points in April 2020. Importantly, exporters more integrated into Global Value Chains (GVCs) and importing higher value emerged as pivotal in bolstering exporter resilience during the crisis, emphasizing the need for policies supporting varied import networks, as well as international trade facilitation. Methodologically, this research innovates by utilizing causal Machine Learning (ML) tools in scenarios where the pervasive nature of the shock hinders the identification of a control group unaffected by the shock, as well as the ex-ante definition of the intensity of the shock's exposure of each unit, making a traditional control group identification unfeasible. This approach effectively predicts firms' trade and uses these predictions to reconstruct the counterfactual distribution of firms' trade under different scenarios and to study treatment effect heterogeneity.

Exports' Survival in New Markets: A firm-level export recommendation model. This paper investigates the factors that better predict a firm's trade status of exporters after expanding to a new destination, specifically whether they continue exporting after two years. Using Colombian customs data, I show that market-level information is crucial for understanding export survival rates, beyond traditional firm-level characteristics like export experience. The paper introduces a novel Machine Learning-based export market entry recommendation tool, designed at the firm-product market level. While firms in the sample did not have access to this tool, the analysis observes which firms chose new destinations that align with the recommendations generated by the tool. Simulated back-testing indicates that firms selecting destinations consistent with the tool's guidance would have experienced a 5 percentage point higher survival rate compared to those choosing other destinations. Additionally, product growth would have been 34 percentage points higher for products where at least one firm followed the tool's suggested market entry, compared to those that did not. The findings suggest that incomplete market insights may lead to sub-optimal export decisions and that exporters incur temporary trade as a way of experimentation and to resolve incomplete market information.

Introduction

The aim of this thesis is to evaluate the impact of different events or policies almost in real time, that work as a quasi-natural experiment, and that have important implications for both the economy and society. The covered events in these essays include the implications of the announcement to withdraw from a political and monetary union (Brexit referendum) on the British economy, the heterogeneous impact of economy-wide shocks (COVID-19) on export dynamics, and the identification of important factors predicting exporter survival. Moreover, this thesis designs a market export recommendation system at firm-product level and evaluates its consequences on export survival and national product growth.

Following the financial crisis of 2008, political parties advocating for departure from the European Union and/or the Euro garnered increasing attention from voters across numerous countries. Party for Freedom in the Netherlands, National Front in France, Five Star Movement and Lega Nord in Italy, and UK Independence Party in the United Kingdom are some examples of populist groups that have based their campaigns on messages against the European club, blamed of being the main obstacle to welfare enhancing policies by limiting national sovereignty. The European Union membership referendum, celebrated in UK on June 23 2016 and known as "Brexit referendum", has been the first success of one of these parties. Both the decision to celebrate the referendum and its result were largely unexpected. While numerous studies in the literature provide forecasts of the cost of leaving the EU for the UK (Dhingra et al., 2017, Ciuriak et al., 2015, Latorre et al., 2019, Ortiz and Latorre, 2018, Kierzenkowski et al., 2016, and citevandenbussche2022global), the economic literature has largely overlooked the anticipatory effects of such a shock on the British economy, including its economic impact prior to the actual departure.

Another unforeseen major event with economy-wide consequences is the COVID-19 pandemic. The outbreak has had a profound impact on the global economy, generating unprecedented health, human, and economic crises. To face the health crisis, governments implemented social distancing and lockdown policies, exacerbating supply and demand shocks (World Bank, 2020). In a highly interconnected world, the impact of the pandemic on international trade has gained great attention (Felbermayr and Görg, 2020; Antràs, Redding, and Rossi-Hansberg, 2023; Bonadio et al., 2020; Evenett, 2020). Global trade, which is typically more volatile than output and tends to fall sharply during a crisis, has shown the biggest fall since the 2009 Global Financial Crisis. From the beginning of the COVID-19 epidemic, scholars underlined that, though its impact on international trade could have been comparable to the Great Trade Collapse of 2008-2009, this time, the demand side shock is accompanied by a supply-side shock (Baldwin and Tomiura, 2020). Moreover, this supply-side effect could be reinforced by a supply-side contagion via importing/supply chains, which have grown in relevance during the last decade (Antràs and Chor, 2022). In other words, supply disruptions in the countries providing intermediate inputs to a given country are likely to hurt also its export performance (Halpern, Koren, and Szeidl, 2015; Navas, Serti, and Tomasi, 2020). All these papers base their identification strategy of the average COVID-19 effect on the cross-country differences in the implementation of lockdown measures over time and study treatment effect heterogeneity by focusing on subsamples or interacting the treatment variable with other possible determinants of heterogeneity. However, the literature disregards the estimate of the causal effect at exporter level (Individual Treatment Effect), so they are not able to classify exporters into most and least affected to estimate the heterogeneity of the analysis, obtaining a narrower image of differential effects from an economy-wide shock in the economy.

Lastly, but closely connected to trade disruptions, Export Promotion Activities (EPA) play a crucial role in facilitating international trade. EPA could be a source to ameliorate the negative effects from the mentioned shocks to the economy, and can enhance poverty reduction and sustainable growth (see Van Biesebroeck, Konings, and Volpe Martincus, 2016). Along these lines, the literature estimates positive and significant effects from EPA to enter new markets (Volpe Martincus and Carballo, 2008; Volpe Martincus and Carballo, 2010; Broocks and Van Biesebroeck, 2017), to increase the GDP per capita (Olarreaga, Sperlich, and Trachsel, 2020), and to boost total exports (Munch and Schaur, 2018). While some EPA initiatives focus on providing financial aid, such as subsidies and credit lines, to promote exports, others prioritize the dissemination of information. For instance, one standard method of information dissemination is through the organization of trade fairs, which enable the establishment of relationships with importers or buyers, sharing information to overcome trade frictions such as product standards, or to learn about foreign demand. The primary objectives of these policies are to reduce the fixed costs associated with acquiring information, facilitate the entry of domestic firms into export markets, and assist exporters in exploring new markets or products. However, budget allocations for export promotion can be insufficient (Olarreaga, Sperlich, and Trachsel, 2020). An easy manner to assist exporters in choosing new markets is to aggregate previous export successful experiences to issue an export recommendations (Hidalgo et al., 2007, Che, 2020, and the International Trade Center (Decreux and Spies, 2016)). However, the literature has largely overlooked the design and impact evaluation of export recommendation policies designed at exporter-product level.

To address these research questions, this thesis adopts a data-driven approach to gain a deeper understanding into the implications of the important policies mentioned earlier. In pursuit of this objective, this work employs both existing and novel methodologies and introduces an export recommendation system at firm-product-market level. The research presented herein endeavors to enrich public discourse with empirical evidence and thoughtful analysis, thereby fostering a more informed and balanced understanding of the impact of various policies and exporter dynamics in developing countries across diverse contexts.

The main challenge in these analyses lies in addressing the *Fundamental Problem of Causal Inference*, as described by (Holland, 1986), for every unit, we fail to observe the value that the outcome would have taken if the chosen level of the treatment had been different or nonexistent. Often, the standard economic literature compares the outcome of treatment units with those from similar control units using techniques such as Differences-in-Differences (Card and Krueger, 1993), with respect to a weighted average of not so similar countries using a synthetic control unit (Abadie, Diamond, and Hainmueller, 2010), or with respect to non-treated units with a similar propensity to be treated (Rosenbaum and Rubin, 1983). However, in the context of such events, selecting the appropriate methodology is not always straightforward or sufficient to accurately estimate the causal impact of the shock. Therefore, it is crucial to identify an appropriate counterfactual to estimate the causal impact of a policy.

The first paper of this thesis delves into the impact of Brexit referendum consequences on the British economy during the period 2016-2017, preceding the formal implementation Brexit. This paper was written together with Francesco Serti under the title "The economic cost of a referendum. The case of Brexit".

Addressing the challenge of not having an economy similar to the British one to estimate the causal effect of this policy, this paper used Synthetic Control Methods (SCM) in a novel manner in this type of literature. By using SCM, the paper estimates a substantial and statistically significant negative effect of the Brexit referendum on the UK's GDP. The study underscores how a policy can exert an impact even years before its implementation. In particular, the paper finds a progressive loss over time, amounting to 1.71% of the observed GDP of the UK in 2017, $Q4$.

The second paper of this thesis introduces a novel methodology to estimate the individual treatment effect and heterogeneity, when no control units are available, resulting from an economy-wide shock on export dynamics in Colombia during the first COVID year, 2020. This paper is co-authored with Marco Dueñas, Massimo Riccaboni and Francesco Serti under the title "Assessing the Impact of COVID-19 on Trade: a Machine Learning Counterfactual Analysis".

This paper shares with the literature the ambition to estimate the causal impact of COVID-19 on trade and its possible heterogeneity. However, this paper uses a different approach based on constructing a counterfactual using the predictive power of Machine Learning to estimate the treatment effect at exporter level (individual treatment effect), that recognises that all firms are directly or indirectly affected by this economy-wide shock. In particular, this study investigates the effectiveness of various Machine Learning (ML) techniques in predicting firms' trade and uses these predictions to reconstruct the counterfactual distribution of firms' trade under different COVID-19 scenarios and to study treatment effect heterogeneity. This work focuses on the probability of Colombian firms surviving in the export market under two different scenarios: a COVID-19 setting and a non-COVID-19 counterfactual situation.

Methodologically, this research innovates by utilizing causal Machine Learning (ML) tools in scenarios where the pervasive nature of the shock hinders the identification of a control group unaffected by the shock, as well as the ex-ante definition of the intensity of the shock's exposure of each unit, making a traditional control group identification unfeasible. This approach effectively predicts firms' trade and uses these predictions to reconstruct the counterfactual distribution of firms' trade under different scenarios and to study treatment effect heterogeneity.

The main finding of this paper reveals that the COVID-19 shock decreased a firm's probability of surviving in the export market by about 20 percentage points in April 2020. Moreover, this study estimates the treatment effect heterogeneity by employing a classification analysis that compares the characteristics of the firms on the tails of the estimated distribution of the individual treatment effects. In particular, exporters more integrated into Global Value Chains (GVCs) and importing higher value emerged as pivotal in bolstering exporter resilience during the crisis, emphasizing the need for policies supporting varied import networks, as well as international trade facilitation.

The third paper makes a significant contribution to the literature by introducing a precise export recommendation model tailored to the firm-product-market level and assessing its impact on export survival rates and national product growth. Furthermore, it identifies important factors for predicting export survival that extend beyond conventional firm performance metrics. Titled "Exports' Survival in New Markets: A firm-level export recommendation model," this work stands as a solo-authored paper.

Moreover, this essay investigates the factors that better predict a firm's trade status of exporters after expanding to a new destination, specifically whether they continue exporting after two years. Using Colombian customs data, the paper finds that market-level information, such as the number of competitors in a given product or market, is critical to understand export survival rates, beyond traditional firm-level factors, such as exporter experience. This information is particularly relevant when predicting firm survival probability in a given market or in a given product-market pair.

In particular, this paper introduces an accurate export market entry recommendation tool powered by Machine Learning and designed at the firm-product market level. Back-testing exercises reveal substantial gains from adhering to the recommendation, showcasing 5 percentage points higher survival probabilities for those exporters trying a new destination suggested by the recommendation rather than just trying a new destination. Furthermore, adopting the recommendation is associated with a notable increase in product growth, with a significant 34 percentage point increase for products where at least one firm followed the guidance compared to those neglecting it or not having any recommendation associated to them. Therefore, this paper underscores the possibility to use a firm level export recommendation to promote poverty reduction and sustainable growth.

Furthermore, this paper posits that exporters could encounter decision-making hurdles stemming from incomplete market insights, highlighting the potential limitations of their understanding of new markets. Additionally, the study suggests that exporters incur temporary trade spells as a form of experimentation and to resolve knowledge uncertainty or incomplete market information, consistent with findings in the literature (Békés and Muraközy, 2012, and Esteve-Pérez, 2021).

Policies aimed at enhancing exporters' capacity to discern optimal destination markets by leveraging the collective exporting experiences of diverse firms could prove beneficial. Strategies such as the introduction of an export recommendation tool coupled with support for off-shore activities like trade fairs and information gatherings may yield positive outcomes not only for exporters' survival rates but also for bolstering product growth at the national level. While the export recommendation tool could be made available by both private and public entities due to its minimal cost, the facilitation and coordination of off-shore activities could be effectively led by Export Promotion Agencies. Such initiatives have the potential to empower exporters with better decision-making tools and resources, ultimately contributing to both individual and collective growth in the country's export landscape.

Together, these papers present a holistic approach that employs various methodologies to evaluate diverse policies and introduces novel tools for comprehending the heterogeneity resulting from shocks when traditional counterfactuals are unavailable. Moreover, they contribute to a deeper understanding of exporter dynamics. Ultimately, this thesis introduces an export recommendation methodology that not only enhances export survival rates but also contributes to national-level product growth.

Chapter 1

The Economic Cost of a Referendum. The Case of Brexit

1.1 Introduction

After the financial crisis of 2008, political parties in favor of leaving the European Union and/or the Euro have increasingly taken the attention of many voters in almost every country. Party for Freedom in the Netherlands, National Front in France, Lega Nord and Five Star Movement in Italy and UK Independence Party in the United Kingdom are some examples of populist groups that have based their campaigns on messages against the European club, blamed of being the main obstacle to welfare enhancing policies by limiting national sovereignty. The European Union membership referendum, celebrated in UK on June 23 2016 and known as "Brexit referendum", has been the first success of one of these parties. Both the decision to celebrate the referendum and its result were largely unexpected.

This comparative case study quantifies the consequences that this decisive referendum has produced on the UK economy so far by using Synthetic Controls Method (SCM) (Abadie and Gardeazabal, 2003; Abadie,

[0]This chapter is based on the work 'The Economic Cost of a Referendum. The Case of Brexit' in collaboration with Francesco Serti (see Ortiz-Giménez and Serti (2020)).

Diamond, and Hainmueller, 2010). By using a data-driven process to build the counterfactual, the SCM reduces ambiguities in the choice of the comparison units and makes transparent the comparability of the characteristics of the treated unit with the control units. In particular, the SCM makes explicit the contribution of each non-treated unit to the construction of the counterfactual and does not allow extrapolating outside the support of the control units, as traditional regression methods do. In our study, we exclude from the baseline "donor pool" of countries (i.e., those which could potentially be used to build the baseline counterfactual for UK in case of no Brexit Referendum) all European countries and United States because, given their high level of economic integration with UK, they could be indirectly affected by the Referendum. The contemporaneously written contribution of Born et al., 2019 estimates the short-run effect of the Brexit referendum by using the SCM and finds similar results. However, in the baseline donor pool, the authors include European countries and do not address the issues raised by Doudchenko and Imbens, 2016 and Ferman and Pinto, 2021. Our study advances the analysis by refining the donor pool selection to exclude European countries, thereby avoiding potential bias from common economic shocks and closer integration with the UK. Moreover, following the recommendations of Doudchenko and Imbens, 2016 and Ferman and Pinto, 2021, we control for unobserved heterogeneity fixed in time and we present the pre-treatment fit obtained after de-trending the series.

In the economic literature, many studies provide forecasts of the cost of leaving the EU for the UK. The most common approaches are econometric models (VAR, NiGEM), New Quantitative Trade Models (NQTM) and Computable General Equilibrium (CGE) models. Using a NQTM Dhingra et al., 2017 quantify the 'static' effects of Brexit on average household income, In an 'optimistic' scenario, the authors estimate that UK will experiment a 1.3% fall in average income (or £893 per household). In a 'pessimistic' scenario, characterized by higher increases in trade costs, they estimate that Brexit will lower its average income by 2.6% (£1,773 per household). Some authors use instead CGE models. In particular, Ciuriak et al., 2015, Latorre et al., 2019 and Ortiz and Latorre, 2018 found that the estimated changes in GDP after Brexit with respect to the no Brexit scenario are $[-2.54\%, -0.97\%]$, $[-2.53\%, -1.23\%]$ and $[-1.15\%, -0.5\%]$, respectively. Moreover, institutions like OECD and Her Majesty's Treasury also carried out their own analysis demonstrating even more dramatic results than the ones presented before. On the one hand, Kierzenkowski et al., 2016 suggest that Brexit will produce losses for British economy equivalent to a non-growth of -3.3% of its GDP. They derived this result by using CGE and NiGEM models. On the other hand, the British government department (Treasury, 2016) estimate that the cost of Brexit will range between $[-7.5\%, -3.8\%]$. To make this prediction, they implemented VAR and NiGEM models. Furthermore, the literature includes pioneering studies investigating the short-term implications of Brexit employing a network approach (Vandenbussche, Connell, and Simons, 2022). In this paper, the authors identify sizable losses in value added in the UK economy resulting in a diminished demand and sector-specific reductions in production and employment. The models used in the literature usually require strong structural assumptions (e.g., trade elasticities, ordering the exogeneity between variables, etc) difficult to validate or understand. However, one of the main strengths of SCM is that it does not require such structural assumptions. Moreover, another strength from SCM is its transparency, as researchers can see which countries contribute to the synthetic UK. Therefore, the results from SCM are robust, intuitive and easy to communicate.

Some papers have concentrated on estimating the benefits from being a member of the European Union, therefore implicitly providing evidence on the cost of leaving the union. According to Haskel, Pereira, and Slaughter, 2007, increases in Foreign Direct Investment (FDI) lead to growth in national productivity. Following Dhingra et al., 2016 (page 11), "multinational firms bring in better technological and managerial know-how, which directly raises output in their operations. FDI also stimulates domestic firms to improve – for example, through stronger supply chains and tougher competition". In the same report, the authors claim that higher trade costs with the EU (after Brexit) are likely to reduce FDI inflows. The researchers also found that, for new entrants, the EU membership exerted a positive effect on FDI: they estimate an increase in FDI of about 28% on average. In this line, Campos, Coricelli, and Moretti, 2018 found that, in the context of EU membership, there are significant and substantial net benefits from deep integration in terms of higher per capita GDP and labor productivity. The study concludes that per capita GDP and labor productivity increased in the United Kingdom due to the EU membership.

Though the bulk of the literature predicts sizable losses for UK after Brexit, some researchers, mass media and politicians still defend that Brexit will not have relevant economic costs for the British economy. We find that these costs emerge even before the UK abandons the UE. The cost of this referendum is estimated in $\$2,376.76$ per citizen in the UK during the year 2017. We also find that this effect is increasing in the time and, therefore, it might be a lower bound for the future consequences of Brexit. In section 1.2 and section 1.3 we describe the empirical methodology and the data. In section 1.4 and section 1.5 we present the main results. section 1.6 is devoted to robustness checks and section 1.7 concludes.

1.2 Model

We estimate the impact of the Brexit referendum on the UK GDP using a comparative case study analysis. Following the seminal work of Abadie, Diamond, and Hainmueller, 2010, we construct a counterfactual for the UK GDP, a so-called synthetic UK, as a weighted average of the GDP of other developed and developing economies. The aim is to choose the weights that are able to replicate the actual dynamics of the UK GDP before referendum and to use them to build the counterfactual GDP of the UK in the post-referendum period.

The effect of the treatment (referendum) on the treated (UK) in this model is represented by:

$$\tau_{it} = Y_{it}^T - Y_{it}^C$$

where:
Y_{it}^T = represents the outcome of unit "i" in case of treatment.
Y_{it}^C = represents the outcome of unit "i" in case of no treatment.

In the post-treatment period, for a treated individual, we can observe only Y_{it}^T, our aim is to estimate Y_{it}^C for the post-treatment periods $t \geq T_0$ (where T_0 is the period when the treatment takes place).

Abadie, Diamond, and Hainmueller, 2010 express the potential outcomes through the following model:

$$Y_{it}^T = \delta_t + \alpha_{it} D_{it} + \mathcal{V}_{it}$$

$$Y_{it}^C = \delta_t + \mathcal{V}_{it}$$

$$\mathcal{V}_{it} = \theta_t Z_i + \lambda_t \omega_i + \varepsilon_{it}$$

where,
Z_t = vector of independent variables
λ_t = common factor (unknown)
θ_t = vector of parameters
ω_i = unobservable term (specific for each country)
ε_{it} = transitory shock (zero-mean)
D_{it} = takes value 1 when the country (i) is treated (UK in this case), and 0 otherwise.
$\alpha_{it} D_{it} = \tau_{it}$

In this setting, Y_{it} and Z_{it} are observables for $N+1$ countries ($i = 1$ represents the country that is treated, and $i = 2, ..., N+1$ represent the non-treated countries). Treatment occurs at time $t = T_0$ and $t \in [1, T]$.

We are going to create a counterfactual for the United Kingdom by constructing a weighted average of a donor pool of countries' outcomes. These weights, $W = (w_2, ..., w_{N+1})$, are allocated to each country such that this counterfactual is as similar as possible to the UK GDP before the treatment. They have the restrictions of being non-negative, $w_i \geq 0$ ($i = 2, ..., N+1$), and to sum up to one, $\sum_{i=2}^{N+1} w_i = 1$. Under these restrictions, SCM safeguards against extrapolation outside the support of the control individuals, following the "philosophy" of matching estimators. Abadie, Diamond, and Hainmueller, 2010 show that as long as we choose W^* such that the outcome and the covariates of the synthetic counterfactual coincide with those of the treated individual ($i = 1$) in the pre-treatment period

$$\sum_{i=2}^{N+1} w_i^* Y_{it} = Y_{1t}$$

$$\sum_{i=2}^{N+1} w_i^* Z_i = Z_1,$$

then

$$\hat{\tau_{it}} = Y_{1t} - \sum_{i=2}^{N+1} w_i^* Y_{it} \quad for\ all \quad t \geq T_0$$

is an unbiased estimator of τ_{it}.

In practice, W^* is chosen solving the following minimization problem:

$$min(X_1 - X_C W)'V(X_1 - X_C W)$$
$$s.t. = \begin{cases} w_i \geq 0 \ (i = 2, ..., N+1) \\ \sum_{i=2}^{N+1} w_i = 1 \end{cases} \quad (1.1)$$

where:
$X_1 = (k\,x\,1)$ is a vector of pre-EU referendum (pre-treatment period) economic characteristics of the treated country. This vector X_1 contains all the pre-treatment values of UK GDP and the pre-treatment elements of vector Z for the UK.
$X_C = (k\ x\ N)$ is the matrix containing the same above characteristics for the N possible control countries (donor pool). In other words, each column of matrix X_C contains the variables of X_1 but for each non-treated country.
$V = (k\ x\ k)$ is the diagonal matrix (symmetric and positive semi definite) reflecting the relative importance of the different characteristics or economic predictors included in X.

Therefore, the optimal W^* minimizes the pre-treatment distance between the vector of characteristics of the treated country and the vector of the synthetic control characteristics according to the metric V. Matrix V is chosen among all positive definite and diagonal matrices such that the mean squared prediction error of the outcome variable during the pre-treatment period is minimized (see Abadie and Gardeazabal 2003).[1]

The SCM rests on two basic identification assumptions. The set of pre-treatment variables chosen should not anticipate the effects of the treatment and donor pool countries (countries that can potentially be selected to create the synthetic control) should not be affected, directly or indirectly, by the intervention. The former requirement means that the effect of the treatment has to start at the date (T_0) chosen without anticipation effects. That is, in the case of Brexit, the effect of the EU leaving referendum should be only observable after UK´s referendum took place. For the latter requirement to hold, we do not include countries that are members of the EU in the donor pool, because they are likely to be affected by the intervention. However, there are also countries outside

[1] In the empirical part, we use the Stata module *synth* specifying the *nested* option.

the EU that could be indirectly affected by the intervention and therefore should not be included in the donor pool. We will further discuss, in the next section, why and which are the countries that are not selected as potential units to create a synthetic counterfactual.

Comparative studies are characterized, as all evaluation studies, by the uncertainty related to the ignorance about the ability of the control group to reproduce the counterfactual of the treated unit in the absence of treatment. Moreover, large sample inferential techniques usually does not apply in this type of studies (because of the small number of units in the control group). We follow Abadie, Diamond, and Hainmueller, 2010 by using exact inferential techniques, that do not require a large number of control units. As in permutation tests, we implement the SCM for each country in the donor pool and we study whether the estimated effect of the treatment on the UK is large enough relative to the (exact) distribution of these estimated placebo effects.

This *in space placebo test* is presented in section 1.5 together with an *in time placebo test*. The latter consists in applying the SCM to the UK randomizing the treatment date (T_0).

In section 1.6 we present several robustness checks. First, inspired by the methodologies introduced by Abadie, Diamond, and Hainmueller, 2015 and Campos, Coricelli, and Moretti, 2018, we address the concern that our results could be driven by the selection of countries present in the donor pool due to spill-over effects or to idiosyncratic shocks which could hit them in the post-treatment period.

Ferman and Pinto, 2021 shows that, in a realistic setting of imperfect pre-treatment fit, the SCM estimator might be inconsistent even in situations in which a fixed effects/DID estimator would not, because the former is not able to wash out time constant unobserved heterogeneity. Therefore, also following Doudchenko and Imbens, 2016,[2] in the second robustness check we employ a model using demeaned data. For this purpose we rewrite the notation of the economic characteristics to ($\ddot{X}_1 = X_1 - \bar{X}_1$) and ($\ddot{X}_C = X_C - \bar{X}_C$), where $\bar{X}_j$ contains the pre-treatment average characteristics for each unit. So that, the new minimization problem that solves for the vector of optimal weights W^* is given by:

[2]Doudchenko and Imbens, 2016 point out that an implicit assumption of SCM is ruling out the possibility that the outcome of the treated unit is systematically larger (or lower) than those of other units. They propose to include an intercept parameter in the objective function of the SCM minimization problem.

$$
\begin{aligned}
&min(\ddot{X}_1 - \ddot{X}_C W)' V (\ddot{X}_1 - \ddot{X}_C W) \\
&s.t. = \left\{ \begin{array}{l} w_i \geq 0 \ (i = 2, ..., N+1) \\ \sum_{i=2}^{N+1} w_i = 1 \end{array} \right.
\end{aligned} \tag{1.2}
$$

If time constant unobserved heterogeneity is additively separable, by using the above demeaned variables, we will take it into account.

According to Ferman and Pinto, 2021, in the presence of non-stationary trends, a close to perfect pre-treatment match would not guarantee the asymptotic unbiasedness of the SCM if there is correlation between treatment assignment and common factors beyond the non-stationary trends. In other words, when pre-treatment fit is imperfect, in models that present a combination of $I(1)$ and $I(0)$ common factors, the SCM estimator does not reconstruct well the factors related to $I(0)$ common factors. Under this framework, $I(0)$ common factors should be uncorrelated with the treatment assignment, so that the SCM estimator is asymptotically unbiased. In a final robustness check, as they suggest, we present the pre-treatment fit after eliminating non-stationary trends and we use it as a diagnosis test for the SCM.

1.3 Data

We use quarterly country-level panel data for the period $2013\ Q1 - 2017\ Q4$. The announcement of the Brexit poll took place on 20 February 2016, whereas the referendum pooling day occurred on 23 June 2016 (our treatment date). Therefore, we consider 13 periods before the treatment (Brexit) and 6 periods after the treatment.

As we mentioned before, the synthetic UK is constructed as a weighted average of potential countries. The control units that belong to the donor pool are 12 non-European countries: Australia, Canada, Chile, Japan, South Korea, Mexico, New Zealand and Turkey (member-countries of the OECD), Brazil, India and South Africa (partner-countries of the OECD) and Russia.[3]

[3]We studied whether the potential donor countries suffer from a structural change (break) in their economic growth during the studied years. We found that, according to Perron Test (Perron, 1997), Indonesia presents a structural break in the $2016Q3$. This in-

The outcome variable Y_{it}, is the real Gross Domestic Product (GDP) in 2010 U.S. million dollars. The pre-treatment characteristics used as predictors in X_1 and X_0 are all pre-treatment outcome lags and the pre-treatment average of some selected predictors based on the literature on economic growth (see Abadie and Gardeazabal 2003): secondary school enrollment (percentage gross), foreign direct investment outflows and inflows (as a percentage of GDP), Business Confidence Index (BCI) and population density.[4]

1.4 Results

By applying the synthetic control method technique we find that the best replication of UK characteristics before the Brexit referendum (2016 Q2) is a weighted average of the following four countries: Mexico, Brazil, Korea and India (ordered from the highest weight to the lowest one). Table 1 shows the set of optimal weights that we use to reproduce UK after the referendum in the absence of Brexit poll. This set of weights is going to be used to obtain the synthetic United Kingdom after the treatment date (2016 Q2).

In the first two columns of Table 2 we compare the pre-treatment values of the explanatory variables of the United Kingdom with those of its synthetic counterpart.[5] The values of the diagonal elements of V linked to each predictor, which are reported in column 3, indicate that, once the pre-treatment values of the GDP are controlled for, the other predictors have very low predicting power (this is common in many applications

tuition was confirmed by Chow Test (Chow, 1960), and this break is due to a sharp fall of Indonesia's GDP. Thus, Indonesia is discarded, despite the fact that, when it was included in the donor pool it got an optimal weight of 0 ($W^* = 0$). That means that Indonesia was not relevant when forming the synthetic counterfactual of UK. We also started our study with China and Israel in the donor pool, but we decided to let them outside the sample because they presented problems of missing values in their datasets. We initially excluded all European member states and Switzerland because they maintain a high level of economic relationships with UK and therefore they could be affected by the treatment. For the same reason, we have not included the United States (US). For the rest of non-European countries, we have not found data for this purpose.

[4]In the appendix A.1, we list more precisely all the variables that are included in the study and their sources. We also tried to use a set of additional predictors: the share of GDP accounted for by agriculture, the share of GDP accounted for by manufacturing, investment and final consumption. Nevertheless, the results remained robust and the fit in the pre-treament period did not improve.

[5]The predictors that are not GDP are averaged over the entire pre-referendum period.

Table 1: Synthetic weights for the United Kingdom

Country	Weight (w*)
Australia	0
Canada	0
Chile	0
Japan	0
Korea	0.179
Mexico	0.554
New Zealand	0
Turkey	0
Brazil	0.201
India	0.066
Russia	0
South Africa	0

of SCM). This explains why the weights that we have obtained produce a close to perfect fit for the lagged values of GDP while for the other variables, whose $V - weights$ are close to zero, some differences persist (Abadie, Diamond, and Hainmueller, 2010).

With the information provided by Table 1 we can estimate the evolution of GDP in the United Kingdom in the absence of Brexit referendum (synthetic UK), after the second quarter of 2016 and compare this series with the actual evolution of GDP for UK. The resulting graph is shown in Figure 1.[6] This figure, jointly with Table 2, suggests that there is a combination of countries that reproduce the economic characteristics of UK before the EU referendum. During the first thirteen quarters (2013Q1-2016Q1), UK and synthetic UK grew almost together. It is only after the referendum, when these two series start diverging. SCM interprets the difference between the GDP of UK and its synthetic counterpart, after the polling date, as the effect of the EU referendum.

Therefore, it makes sense to plot the difference between the two series, so that we can more easily evaluate how the GDP of the synthetic unit evolves compared to the one of UK. Figure 2 shows the evolution of the GDP gap between these two units over the time. This gap is the esti-

[6] The "x" axis of Figure 1 and of the other figures represents the time periods in quarters. Being the quarter $1 = 2013Q1$, the quarter $14 = 2016Q2$ (treatment period) and the quarter $20 = 2017Q4$.

Table 2: Balancing of predictors and V-weights

	UK	Synthetic UK	Predictor V-weight
GDP (2013 Q1)	2343303	2343567	0.08
GDP (2013 Q2)	2355970	2353594	0.00
GDP (2013 Q3)	2376004	2377849	0.00
GDP (2013 Q4)	2388334	2394805	0.18
GDP (2014 Q1)	2408849	2412180	0.14
GDP (2014 Q2)	2429396	2431156	0.00
GDP (2014 Q3)	2447887	2445851	0.16
GDP (2014 Q4)	2466521	2468233	0.00
GDP (2015 Q1)	2474971	2477142	0.00
GDP (2015 Q2)	2489198	2488505	0.20
GDP (2015 Q3)	2499567	2505913	0.00
GDP (2015 Q4)	2517709	2510589	0.11
GDP (2016 Q1)	2523036	2531478	0.13
DNPOP	267.4253	162.3938	3.948e-07
SCSEC	125.7761	94.14681	1.330e-07
FDII	52.3811	30.2239	6.918e-07
FDIO	58.22206	12.1585	5.269e-07
BCI	100.9391	99.79369	1.158e-07

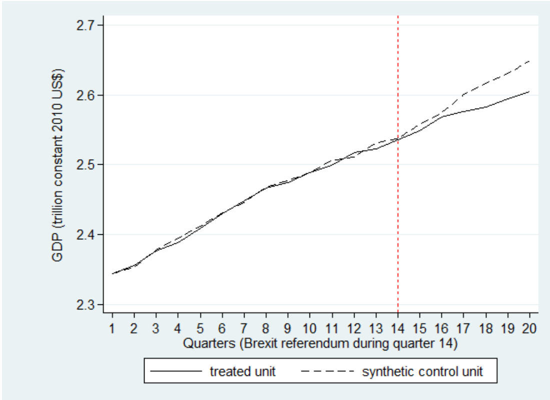
Note: Columns 1-2 of this table present the characteristics of UK and its synthetic control before the Brexit referendum. The last five predictors (population density, secondary school enrollment, foreign direct investment inflows/outflows, business confidence index) are averaged across the entire pre-treatment period. Column 3 presents the optimal V-weight that each predictor receives.

mated difference in GDP as a consequence of the EU referendum for the UK ($\hat{\tau_{it}}$). Despite the fact that the United Kingdom has experienced an increase in real GDP after the EU referendum, Figure 2 suggests that this increase would have been much higher if EU referendum would not have taken place (Governor of the Bank of England, Mark Carney, declared on November 2017 that "British economy should really be booming, and it is just growing").[7] It is worth mentioning that GDP gap is increasing in time, which means that the effect of Brexit referendum might be only a lower bound of the future Brexit effect.

We evaluate the size of Brexit referendum effect compared to the size of British GDP in Table 3. In the third column we report the estimated GDP gap as a percentage of the actual GDP. This ratio is increasing for almost every period. In the last quarter of 2017, GDP in synthetic UK is estimated to be $\$44.5$ billion higher than in the actual United Kingdom.

[7] Carney declared the above statement on 5 November 2017 to Robert Peston, a British journalist, on his TV program "Peston on Sunday".

Figure 1: GDP of UK and its synthetic counterpart.



Note: This figure presents the GDP of the UK and the estimated GDP for the synthetic UK before and after the Brexit referendum.

Hence, Brexit represents a fall (economic cost) of 1.71% over the actual GDP for period 2017 $Q4$. During the whole period (2016Q3-2017Q4) that we analyze, the United Kingdom should have grown $\$156.03$ billion more than it actually grew. This cumulative non-growth represents 1.01% over the total cumulative GDP for the same period of time (six quarters for which we have available data), a cost of $\$2,376.76$ per UK citizen.[8]

1.5 In space and in time Placebo effects

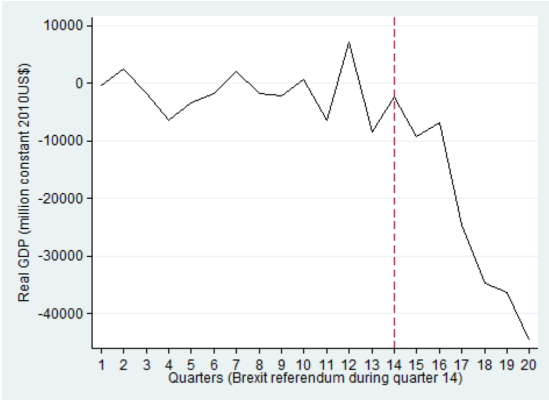
In this section we apply exact inference techniques to "in-space" placebo experiments and we perform an "in-time" placebo study.

The first placebo study is conducted "in space", by assigning the original treatment to each country included in the donor pool.[9] Therefore,

[8]The cost of the Brexit referendum can also be expressed in British pounds by applying the exchange rate as of December 29, 2017, which was 0.7441 GBP/USD. Accordingly, the estimated cost for the British economy is £116.10 billion, with a corresponding cost of £1,768.42 per UK citizen.

[9]Figure 3 does not present the estimated treatment effect for India and New Zealand because the model is not able to reproduce the GDP pre-treatment path of these two coun-

Figure 2: Estimated difference between the actual and the synthetic GDP of UK.



Note: This figure presents the gap between the UK series and its synthetic counterpart. It shows that before Brexit referendum the gap is constant and it is around zero. However, after the poll the gap drops sharply.

Table 3: Size of GDP gap

	Observed GDP (2010 US$ millions)	GDP gap (US$ millions 2010)	Gap over original GDP (%)
2016-Q3	254875.75	-9281.91	-0.36%
2016-Q4	2567829.25	-6873.86	0.27%
2017-Q1	2575888.5	-24594.16	-0.95%
2017-Q2	2582240.75	-34593.79	-1.34%
2017-Q3	2594141.75	-36196.55	-1.40%
2017-Q4	2604230.5	-44489.34	-1.71%

Note: Columns 1 and 2 present the data of the actual value GDP and the estimated GDP gap of the UK, respectively. Column 3 reports the gap as a percentage of the actual GDP.

we assign the treatment to countries that did not experienced it. We observe the effect of the treatment for these countries in Figure 3. The effect obtained in Figure 2 is considerably larger when compared with the placebo effects estimated for the donor countries in Figure 3 (only for

tries. i.e. the optimization routine is not able to find the minimum of equation 1.1. One of the virtues of SCM is exactly that it does not allow extrapolating outside the support of the control units, as traditional regression methods do.

Russia we estimate a negative, though imprecise, effect). Therefore, we consider that the effect of Brexit referendum for the United Kingdom is big enough to be considered significant.

A more refined way of testing the significance of the estimated effect of Brexit is considering the ratio of the of post-$2016Q2$ MSPE (Mean Squared Predictor Error) over the pre-$2016Q2$ MSPE for the United Kingdom and the other control countries. The mean squared predictor error measures the average gap for each country with respect to its synthetic counterpart. A large post-treatment MSPE does not indicate a large effect of the treatment if the synthetic control does not match the real data accurately before the treatment (i.e. if pre-treatment MSPE is also large). For this reason, Abadie, Diamond, and Hainmueller, 2010 suggested to compare the ratio between post-intervention MSPE and pre-intervention MSPE of the treated unit to those of the units contained in the donor pool. Figure 4 shows that for the United Kingdom the post-treatment MSPE is about 50 times larger than the pre-intervention MSPE. The MSPE ratio for UK is well above the MSPE ratios for the control countries. As shown by Abadie, Diamond, and Hainmueller, 2010, we can use these MSPE ratios to test the null hypothesis that the treatment has no effect and it is assigned randomly. The chance to select a country at random, with a MSPE ratio as high as the one for the UK, is $1/11 \simeq 0.09$, that is below the conventional 10% level of significance used in statistics. Then, we reject the null hypothesis and conclude that Brexit referendum had a significant effect on UK GDP.

The second placebo study is conducted "in time", by reallocating the date of the Brexit referendum to a period before the current referendum in order to assess anticipation effects. In particular, we set a new artificial date for Brexit referendum, $2015Q2$ (just in the middle of the time period horizon).[10] Figure 5 reports the results obtained when applying the SCM using the same donor pool but setting $T_0 = 2015Q2$. The synthetic United Kingdom almost perfectly replicates the trend of GDP in the UK along the period $2013Q1 - 2015Q1$ (before treatment). But most importantly, GDP trends, for both the UK and the synthetic UK, do not diverge until the period $2016Q4$, as in the baseline estimates. Until this moment, both units grew together. There is no jump around the new artificial treatment date.

[10] Note that period 10 corresponds to the date 2015Q2.

1.6 Robustness

The first two robustness checks investigate the sensitivity of our results to the chosen weights, i.e. how dependent are our results on the specific control countries. Indeed, our results could be biased by unobserved shocks and treatment spillovers affecting some donor countries in the post-treatment period. The last two robustness checks deal with the concerns raised by Ferman and Pinto, 2021.[11] They suggest removing time-constant unobserved heterogeneity from the series and studying the performance of the SCM after detrending the series.

The first robustness check can be presented as a combination of four different tests inspired by the work of Abadie, Diamond, and Hainmueller, 2015. In all of them we exclude some countries from the donor pool and we study how the estimated effect of Brexit changes when we try to reproduce the synthetic UK using these restricted samples. Figure 6 summarize the results of these analyses. The black dashed line identifies the actual GDP of UK, the gray thick dashed line corresponds the baseline synthetic UK estimated in section 1.3 and the thin gray dashed lines indicate the synthetic UKs obtained by restricting the baseline donor pool. In graph (a), we exclude a pair of countries from the donor pool at random and we repeat this step 10 times. In graph (b), we present the results obtained by discarding a pair of countries among those that received positive weights in the baseline estimates. In this way we obtain other six alternative donor pools.[12] In graph (c), we apply the SCM by using the four possible donor pools obtained by removing one of the countries that received a positive weight in the baseline estimates. Finally, in graph (d) we report the estimates stemming from the four possible donor pools which can be obtained by excluding at the same time a different triplet of countries among those that received a positive weight in the baseline estimates. In all the four graphs, we observe that the synthetic units generated fit accurately the series of GDP for the United Kingdom before 2016Q2. Moreover, after this period, these synthetic units diverge, similarly to what happens to the baseline synthetic UK estimated in section 1.3. Consequently, this figure suggests that the baseline results presented in the previous section are not driven by the inclusion in the donor pool of some specific country, in this way alleviating possible concerns about unobserved post-Brexit shocks such as having Common Wealth

[11] See section 1.2 and the additional explanations given below in this section.

[12] Remember that in the baseline estimates only four countries obtained positive weights.

members among the countries with a positive weight creating the synthetic UK, as it is the case of India.

For the second robustness check we augment the number of potential control countries by allowing other European countries to enter in the donor pool.[13] We take this decision after verifying that we do not have enough evidence showing that the GDP of these European countries is affected by this referendum (i.e. that that there are spillover effects of the referendum).[14] Among the countries belonging to this augmented set, made up by 34 countries (the baseline 12 countries and 22 European countries), we first select randomly 1000 donor pools made up by 12 countries, which is the size of the baseline donor pool. Then we estimate with the SCM the cost of the Brexit referendum for UK using these different and randomly chosen donor pools and we compare these estimates with our baseline estimate. Figure 7 shows the results of this robustness check. The black straight line represents the actual GDP of the United Kingdom, the dashed black line represents the baseline synthetic UK (built using the baseline donor pool made up by 12 non-European countries) and the gray dashed lines represent the different synthetic UKs estimated using these randomly chosen donor pools. Following the logic of Abadie, Diamond, and Hainmueller, 2010, in Figure 7 we present only the synthetic UKs which attain a pre-treatment fit similar to the one obtained by the baseline synthetic UK and are, therefore, comparable with our baseline estimates. As a measure of pre-treatment fit we adopt the pre-treatment normalized mean squared error index introduced by Ferman, Pinto, and Possebom, 2020:[15]

$$\widetilde{R}^2 = 1 - \frac{\frac{1}{T_0-1}\sum_{t=1}^{T_0-1}(Y_{1t}-\hat{Y}_{1t})^2}{\frac{1}{T_0-1}\sum_{t=1}^{T_0-1}(Y_{1t}-\bar{Y}_1)^2}. \tag{1.3}$$

[13]The countries included are Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland. We do not consider Iceland and Ireland because they do not have data for the Business Confidence Index (BCI) variable.

[14]We have estimated the synthetic unit for each of these EU countries by using the baseline donor pool of countries (non-European states). We do not find any significant effect of the Brexit referendum on the GDP of each of these economies (see Appendix A.2).

[15]This measure will be equal to 1 when the pre-treatment fit is perfect. However, differently from the standard R^2, it can be negative.

where Y is the observed GDP of the real UK, $\hat{Y}$ is the estimated GDP of the synthetic UK and $\bar{Y}_1 = (\sum_{t=1}^{T_0-1} Y_{1t})/(T_0 - 1)$.

Given that the $\widetilde{R}^2$ of the baseline synthetic UK is about 0.99, we retain in Figure 7 those synthetic UKs with a $\widetilde{R}^2 \geq 0.9$ (i.e. about 70% of them; the results are not sensitive to changing this threshold). We find that the bulk of these synthetic UKs, built by randomly choosing different donor pools made up by 12 countries, imply similar or greater estimated losses due to the Brexit referendum. Therefore, also according to this robustness check, our results are not sensitive to changes in the countries included in the donor pool.[16]

The third robustness check is presented in Figure 8. Following equation 1.2, before applying the standard SCM we purge its average pre-Brexit GDP from every country's GDP.[17] As explained in section 1.2, in the realistic setting of imperfect pre-treatment fit, and differently from the standard SCM, this demeaned SCM is robust to time constant unobserved heterogeneity correlated with selection into treatment. As shown in in Figure 8, the pre-treatment fit obtained using this demeaned SCM is close to perfect and the estimated treatment effect is very similar to the baseline result.

Finally, following the suggestions of Ferman and Pinto, 2021, we check whether the close-to-perfect pre-treatment fit obtained for the GDP non-stationary series hides discrepancies in common factors beyond these non-stationary trends. Indeed, when such residual (and stationary) common trends are correlated to treatment assignment they may lead to asymptotic bias. By following the approach of Ferman and Pinto, 2021, firstly we try to wash out the non-stationary common factor by subtracting, to controls and treated countries, the average of the controls' GDP in each period of time. If a common factor that has the same effect for every unit is the reason behind the non-stationarity, the resulting series should not feature non-stationary trends. Secondly, as an alternative procedure to

[16]This robustness check is probably less accurate than the previous one, because we cannot be completely sure that there are no spillover effects on the aggregate of the rest of European countries. Indeed, it might be the case that the poll affects these countries as a group, even though this effect is not significant for each European country.

[17]For this and the last robustness check, we use the baseline set of donor countries (i.e. we do not consider European countries). Moreover, following Ferman and Pinto, 2021, we exclude from the explanatory variables all the variables different from GDP. Results including the same set of explanatory variables of the baseline specification are basically identical.

de-trend the data, we subtract the trend obtained by fitting a fourth order polynomial to the GDP series. As shown in Figure 9 and Figure 10, both ways to de-trend the data seem successful: the series for the UK and the synthetic UK do not present a clear non-stationary trend. The pre-treatment fit obtained after eliminating the non-stationarity ($\widetilde{R}^2 = 0.65$ and $\widetilde{R}^2 = 0.55$ for the first and second de-trending procedures, respectively) is worst than the one obtained using the baseline not de-trended estimates (i.e. $\widetilde{R}^2 = 0.99$). This reduction in pre-treatment fit obtained after eliminating non-stationary trends is in line with the examples reported in Ferman and Pinto, 2021, which apply this diagnosis tool to the influential studies of Abadie and Gardeazabal, 2003 and Abadie, Diamond, and Hainmueller, 2015. Moreover, it is also reassuring to notice that the pre-treatment gap between the real and the synthetic UK is far less relevant than the one observed during the post-treatment period.

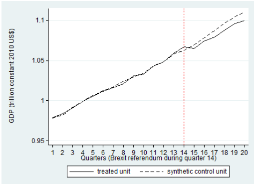
1.7 Concluding discussion

This paper estimates the effect that the European Union membership referendum has had on the British economy so far, focusing on its consequences on GDP. Using the SCM (Abadie, Diamond, and Hainmueller, 2010), we construct the counterfactual GDP of the United Kingdom, after June 2016, as if the British referendum had not taken place. This paper quantifies the cost of the European Union membership referendum for the British economy in $\$156.03$ billion,[18] since the period when the referendum took place until December 2017. This figure is the equivalent to an accumulated non-increase in GDP of 1.01 percent. Our study finds that the referendum caused a lower GDP growth than the one that the UK would have experienced in the absence of referendum during almost all the periods analyzed after the Brexit poll. Furthermore, we found that this non-growth effect is increasing in time.

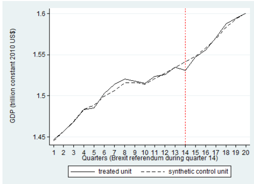
The evidence found in this paper complements the analyses predicting the future consequences of Brexit for UK by demonstrating that the outcome of the referendum itself has already produced a slowdown of the UK economy, before the actual Brexit has taken place. This finding is an early warning for populist European parties against irresponsible and unrealistic plans of abandoning smoothly the EU common market.

[18] The cost of the Brexit referendum can also be expressed in British pounds by applying the exchange rate as of December 29, 2017, which was 0.7441 GBP/USD. Accordingly, the estimated cost for the British economy is £116.10 billion.

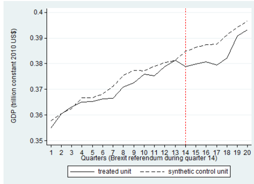
Figure 3: In-space Placebo Test.



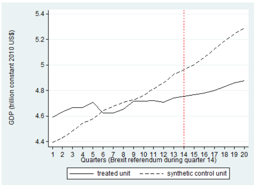
Synthetic Australia



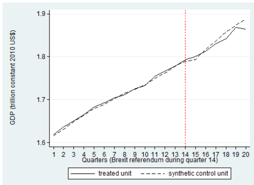
Synthetic Canada



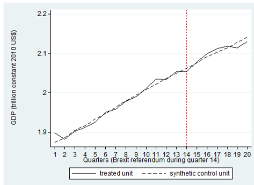
Synthetic Chile



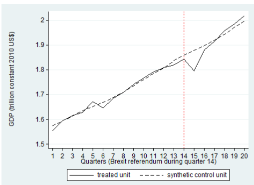
Synthetic Japan



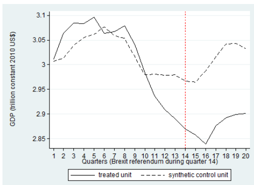
Synthetic Korea



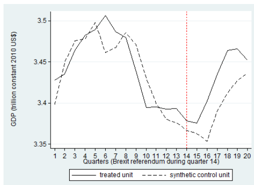
Synthetic Mexico



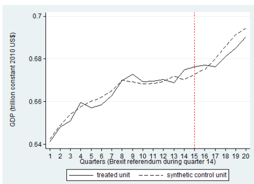
Synthetic Turkey



Synthetic Brazil



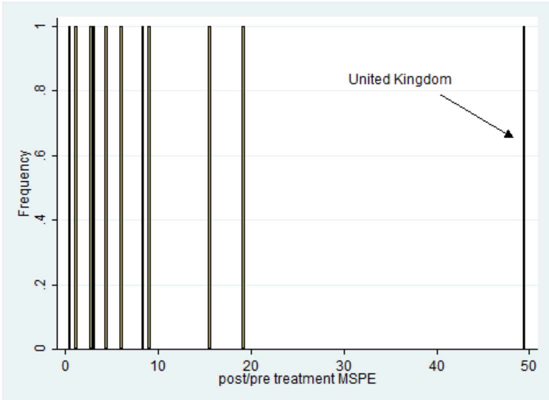
Synthetic Russia



Synthetic South Africa

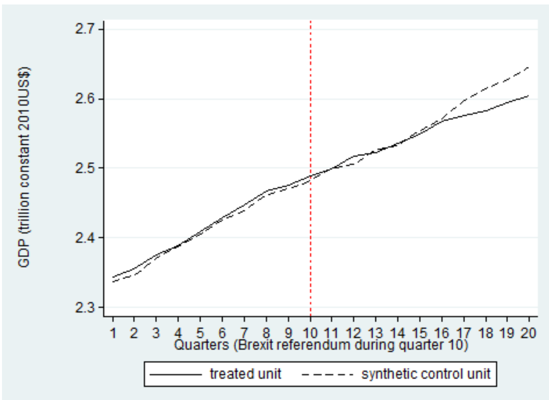
Note: We assign the treatment to countries (used in the donor pool) that have not experienced it and we estimate its (placebo) effect by SCM. The effect for UK reported in Figure 2 is considerably larger when compared with the effects for the donor countries presented in this figure.

Figure 4: Ratio of post Brexit referendum MSPE over pre Brexit referendum MSPE for the UK and the donor countries.



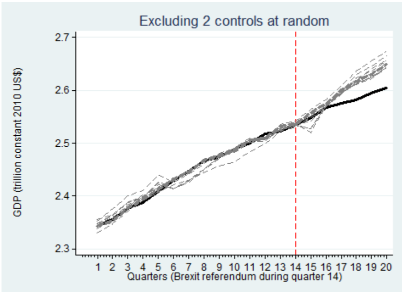
Note: This figure presents the ratios of post/pre MSPE (Mean Squared Predictor Error) for the UK and the countries in the donor pool.

Figure 5: In-time Placebo Test: GDP of the United kingdom and GDP of the Synthetic United Kingdom.

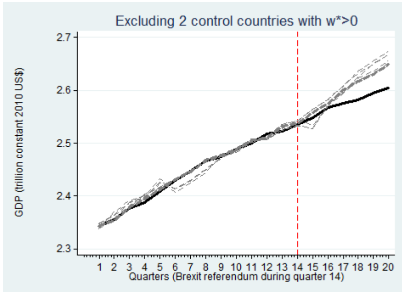


Note: This graph shows the GDP of the UK and its synthetic counterpart obtained by imposing a different period of treatment ($T_0 = 2015Q2$).

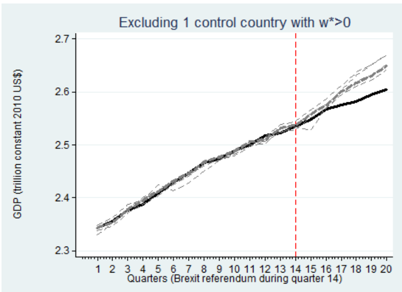
Figure 6: Robustness check using the original donor pool sample.



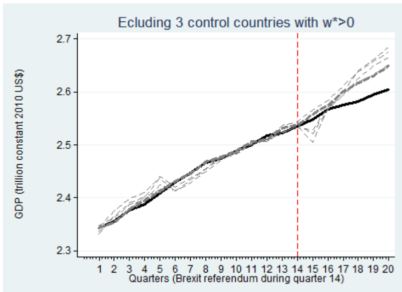
(a)



(b)



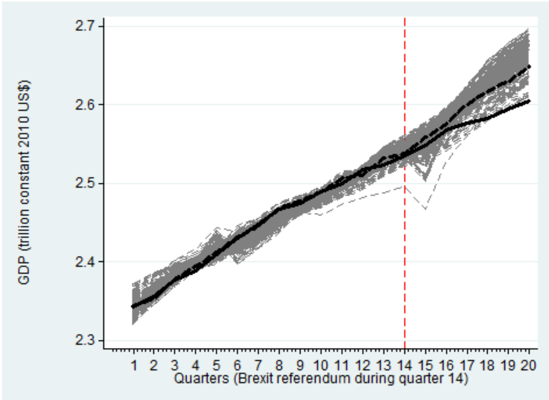
(c)



(d)

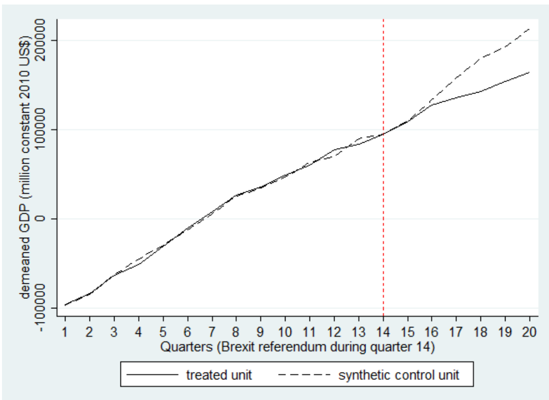
Note: In this robustness check, we restrict the original donor pool of countries in 4 different ways, in order to study how robust are the results to changes in the composition of the pool. In Figure a) we extract, by pairs, countries from the baseline donor pool (at random). In Figure b) we extract, by pairs, countries from the baseline donor pool that have received $w^* > 0$ in the baseline estimation. In Figure c) we extract, individually, countries from the baseline donor pool that have received $w^* > 0$ in the baseline estimation. In Figure d) we extract from the donor pool groups of three countries that have received $w^* > 0$ in the baseline estimation.

Figure 7: Random donor pools.



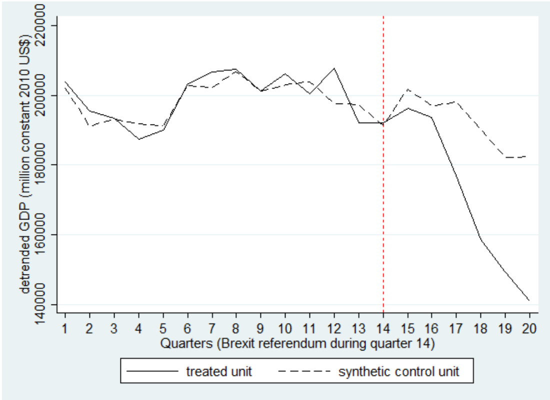
Note: The black line of this figure represents the GDP of the UK and the back dashed line its baseline synthetic GDP. Each gray dashed line represents the synthetic GDP of UK estimated with randomly chosen donor pools potentially containing both European and Non-European countries. Each random donor pool contains 12 countries as the baseline donor pool and attains an $\widetilde{R}^2 \geq 0.9$.

Figure 8: Demeaned SCM.



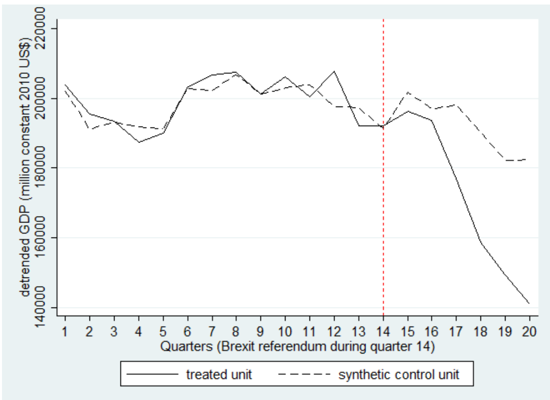
Note: This figure present the results obtained by using the demeaned GDP following equation 1.2.

Figure 9: De-trended output by extracting the pre-Brexit average GDP of the controls to the treated and control units.



Note: This figure presents the de-trended series of GDP for both the UK and the synthetic UK. We have de-trended the series by extracting, to treated and control countries, the average of the controls' output in every period of time.

Figure 10: De-trended output by fitting a fourth order polynomial.



Note: This figure presents the de-trended series of GDP for both the UK and the synthetic UK. We have de-trended the series by fitting a fourth order polynomial.

Chapter 2

Assessing the Heterogeneous Impact of Economy-Wide Shocks: A Machine Learning Approach Applied to Colombian Firms.

2.1 Introduction

The COVID-19 outbreak has affected the world economy, generating unprecedented health, human, and economic crises. To face the health crisis, governments implemented social distancing and lockdown policies, exacerbating supply and demand shocks (World Bank, 2020). In a highly interconnected world, the impact of the pandemic on international trade has gained great attention (Felbermayr and Görg, 2020; Antràs, Redding, and Rossi-Hansberg, 2023; Bonadio et al., 2020; Evenett, 2020).

[0] This chapter is based on the work ’Assessing the Impact of COVID-19 on Trade: a Machine Learning Counterfactual Analysis’ in collaboration with Marco Dueñas, Massimo Riccaboni and Francesco Serti (see Dueñas et al. (2021)).

Global trade, which is typically more volatile than output and tends to fall sharply during a crisis, has shown the biggest fall since the 2009 global financial crisis. From the beginning of the COVID-19 epidemic, scholars underlined that, though its impact on international trade could have been comparable to the Great Trade Collapse of 2008-2009, this time, the demand side shock is accompanied by a supply-side shock (Baldwin and Tomiura, 2020). Moreover, this supply-side effect could be reinforced by a supply-side contagion via importing/supply chains, which have grown in relevance during the last decade (Antràs and Chor, 2022). In other words, supply disruptions in the countries providing intermediate inputs to a given country are likely to hurt also its export performance (Halpern, Koren, and Szeidl, 2015; Navas, Serti, and Tomasi, 2020).

This paper examines the heterogeneous effects of an economy-wide shock, specifically COVID-19, on export dynamics, with a focus on firm-level survival in Colombia during 2020. We aim to identify the causal impact of the pandemic on firms' likelihood of maintaining their presence in export markets. Our methodological contribution lies in advancing causal Machine Learning (ML) techniques to address the challenges posed by the pervasive nature of an economy-wide shock, which complicates the identification of a traditional control group and the ex-ante determination of exposure intensity. By employing these ML tools, we generate accurate predictions of firms' trade outcomes, enabling the reconstruction of counterfactual trade distributions under various scenarios and facilitating the analysis of treatment effect heterogeneity. We find that that the COVID-19 shock reduced a firm's probability of survival in the export market by approximately 20 percentage points in April 2020. However, this average effect masks substantial heterogeneity across and within sectors. Specifically, we observe that smaller firms and those with greater exposure to severely affected export destinations and import sources are disproportionately impacted by the pandemic's containment measures.

Our analysis reveals a critical insight, the degree of geographical diversification and involvement in the import market play a pivotal role in determining the resilience to the COVID-19 shock. This highlights the strategic advantage held by firms that possess diversified sourcing locations and engage in import activities. Integration into Global Value Chains (GVCs) not only safeguards against the disruptive effects of crises, it also shows how a globally connected and diversified approach supports export resilience. This evidence points to the importance of supportive policies like promoting trade agreements, improving infrastructure, and making regulatory reforms to facilitate trade. These actions could significantly strengthen sectors that depend on exports, making them more stable and adaptable during uncertain times, ultimately boosting economic progress and sustainable growth. Understanding the heterogeneous effects of economy-wide shocks is a crucial undertaking as it represents the foundational stage in devising policy interventions intended to mitigate their negative outcomes and reactivate economic activities.

Our paper contributes to a growing body of literature that applies Machine Learning (ML) techniques in the field of international trade. (Breinlich et al., 2022) employ ML algorithms to analyze the impact of a large number of trade agreement provisions on trade flows; Che, 2020 uses ML to provide product recommendations to optimize a country's export structure. Distinctively, our study is the first to employ ML techniques to analyze firm-level survival in international trade, providing novel insights into the determinants of firm performance, allowing to assess the heterogeneous impact of economy-wide shocks in the absence of a control group unaffected by the shock.

Moreover, this paper contributes to the literature using ML to estimate counterfactuals (Cerqua and Letta, 2020; Fabra, Lacuesta, and Souza, 2022). In this literature it is common to estimate causal effects by comparing the counterfactual predictions with the observed outcome in case of treatment, following the so-called "consistency assumption": if the outcome in case of treatment is observed then it also represents the potential outcome under treatment. Instead, we follow Chernozhukov et al., 2020 by using ML techniques to reconstruct firm potential outcomes in case of no treatment and to predict the outcomes in the treatment scenario. Our findings emphasize the importance of comparing estimated counterfactual outcomes with predictions for the treatment scenario (rather than directly with observed outcomes during treatment). This comparison helps rectify estimation errors resulting from imperfectly reconstructing unseen counterfactuals, particularly when the objective is to determine the factors contributing to treatment effect heterogeneity, and not just to estimate the Average Treatment Effect.

Methodologically, this study represents a pioneering adaptation and application of the generic machine learning framework proposed by Chernozhukov et al., 2020 to contexts where a contemporaneous control group is unavailable. While traditional settings, such as randomized controlled trials (e.g., Atkin, Khandelwal, and Osman, 2017), rely on the clear identification of unaffected groups, our approach modifies these tools to estimate treatment effects in environments where identifying ex-ante unaffected firms is challenging. This innovation not only extends the applicability of these machine learning techniques but also addresses a crucial gap in empirical methods for analyzing widespread, systemic shocks.

This paper is connected to the literature on the heterogeneous impact of the COVID-19 shock on trade, De Lucio et al., 2022, Berthou and Stumpner, 2022, Espitia et al., 2021, and Liu, Ornelas, and Shi, 2021. All these papers base their identification strategy of the average COVID-19 effect on the cross-country differences in the implementation of lockdown measures over time and study treatment effect heterogeneity by focusing on subsamples or interacting the treatment variable with other possible determinants of heterogeneity. However, we use a different approach that assumes that all firms are directly or indirectly affected by this economy-wide shock and that it is very challenging to define ex-ante a variable summarising the (differential) exposure of firms to COVID-19. Additionally, our method for estimating the heterogeneity analysis safeguards against the risks of overfitting and multiple testing. Among the possible determinants of heterogeneity, we also consider a firm's diversification on the export and import side. Therefore, our study is also related to the international trade literature on the role of diversification in mediating the impact of adverse shocks (Kramarz, Martin, and Mejean, 2020; Grossman, Helpman, and Lhuillier, 2023; Lafrogne-Joussier, Martin, and Mejean, 2022).

The remainder of the paper is structured as follows. Section 2.2 details our empirical strategy. Section 2.3 presents the firm-level data, variables employed in the analysis, and descriptive statistics. Section 2.4 reports the estimation results, and Section 2.5 summarizes our findings and discusses the relevance and limitations of our analysis.

2.2 Methodological framework

This section lays out our empirical approach to estimating the effect of an economy-wide shock on firms' survival probabilities in export markets and exploring its heterogeneity based on firms' observable attributes. Moreover, in this section, we introduce a methodological advancement that extends the framework developed by Chernozhukov et al., 2020 to address situations where a contemporaneous control group cannot be readily identified. Our approach adapts the generic machine learning technique to estimate individual treatment effects in contexts where the shock or intervention affects a broad set of firms, making it difficult to distinguish unaffected firms ex-ante. This modification enhances the robustness of causal inference in settings characterized by widespread or systemic shocks, offering a valuable tool for empirical research in environments with such constraints.

As with any evaluation study, the primary identification task is to find a counterfactual outcome for the affected units, which is unobserved. In the presence of an economy-wide shock such as COVID-19, no subset of unaffected firms can likely be selected as a control group because the shock impacts, at least indirectly, all firms. Furthermore, even an identification strategy based on comparing individual firms subject to different treatment intensities appears infeasible due to the complex and ex-ante unknown paths through which firms are potentially exposed to the treatment. Though we study whether treatment effect heterogeneity depends, inter alia, on measures of the exposure to the COVID-19 shock,[1] the intensity of treatment might also depend on other firms' characteristics, such as the identity of suppliers and clients, the characteristics of the traded final product, among many others, that we cannot know in advance and whose interactions are *a priori* unknown.[2]

Therefore, as it is often practised when studying large-scale shocks' effects, we must employ the information on the pre-shock behavior to estimate the counterfactual behavior (in the hypothetical situation of the absence of the shock) during the actual shock. This process involves forecasting the future conduct of entities based on their historical behavior, an application perfectly suited to ML techniques, which are designed for such out-of-sample prediction tasks.

In line with the reasoning of Varian, 2016, and drawing parallels with the applications employed by Cerqua and Letta, 2020, and Fabra, Lacuesta, and Souza, 2022, we harness the predictive strength of ML techniques. This allows us to construct a hypothetical scenario for firm-level outcomes during the shock period, using pre-shock data concerning firms' export behaviors and attributes. We aim to study whether a firm that was exporting in a given month during the pre-shock year t_{s-1} (that is 2019 in our specific application) will export again during the same month of the

[1]These indexes are described in detail in section 2.3. They are based on firms' past export and import activities in different countries and on the time-varying strength of the virus and the stringency of the policies aimed at mitigating its spread.

[2]A possible strategy to identify a control group would rely on firm's statements about the consequences of the COVID-19 crisis on their activities.

year of the shock t_s (that is 2020 in our specific application). The empirical analysis we shall describe is carried out for each month separately to allow the effects of the explanatory variables (e.g., the hypothesized determinants of firm export status) to vary throughout the year.[3]

Let us define the potential outcome under the scenario $d \in \{0, 1\}$ for a firm i at the time t as Y_{it}^d. In this case, d represents an indicator variable signifying the presence of the shock. More precisely, the outcome of our interest is the firm's export status, which is equal to one if a firm exporting at year $t-1$ in a given month is also exporting at year t during the same month (otherwise it will be zero). Note that the regressors, X_{it}^d, may also be influenced by the presence or absence of the shock.

The initial step in our analysis involves estimating the counterfactual outcome at t_s: Y_{i,t_s}^0. The model we utilize to derive this counterfactual (and the counterfactual itself) is referred to as the "Shock Unaware Machine" (SUM), a term acknowledging the ML techniques employed in constructing the counterfactual and the fact that no information about the shock is used in the analyses. In particular, we use the outcomes and covariates observed at $t_s - 1$ and $t_s - 2$ to reconstruct $Y_{t_s}^0$ based on the following assumptions (for ease of notation, we shall omit the subscript i from this point forward):

1. Neither covariates nor outcomes of $t_s - 2$ and $t_s - 1$ are affected by the shock:

$$Y_t = Y_t^0 = Y_t^1, \quad X_t = X_t^0 = X_t^1 \quad \text{for} \quad t = (t_s - 1, t_s - 2) \tag{2.1}$$

2. Define $Y_t^0 = f_t^0(X_{t-1}^0) + u_t^0$, where $f_t^0(\cdot)$ is a generic model or function representing the relationship between explanatory variables and the outcome in the absence of the shock such that $\mathbf{E}[Y_t^0 | X_{t-1}^0] = f_t^0(X_{t-1}^0)$. Under (i), for $t = t_s - 1$ we have that $Y_{t_s-1} = f_{t_s-1}^0(X_{t_s-2}) + u_{t_s-1}^0$ such that $\mathbf{E}[Y_{t_s-1} | X_{t_s-2}] = f_{t_s-1}^0(X_{t_s-2})$. The second assumption states that the function f_t^0 does not depend on t, i.e., it is stable over the two considered years:

$$f_{t_s-1}^0 = f_{t_s}^0 = f^0. \tag{2.2}$$

Therefore, under the above assumptions, we can write $Y_{t_s}^0 = f^0(X_{t_s-1}) + u_{t_s}^0$, such that $\mathbf{E}[Y_{t_s}^0 | X_{t_s-1}] = f^0(X_{t_s-1})$, and we can use data on $t_s - 2$ and

[3] A yearly model with month fixed effects would just allow for different levels of the probability of exporting for each month, e.g. a different constant for each month.

$t_s - 1$ to estimate $Y^0_{t_s-1} = f^0(X_{t_s-2}) + u^0_{t_s-1}$ and retrieve $\hat{f}^0$. By applying this invariant estimated function to the covariates of $t_s - 1$, we can obtain the predictions for the counterfactual (without the shock) outcome in t_s:

$$\hat{Y}^0_{t_s} = \hat{f}^0(X_{t_s-1}) = Y^0_{t_s} - \overbrace{\mathcal{E}^0_{t_s}(X_{t_s-1})}^{\textit{Prediction error}} - \overbrace{u^0_{t_s}}^{\textit{Orthogonal error}} . \tag{2.3}$$

In general, the estimated counterfactual outcome at time t_s, denoted by $\hat{Y}^0_{t_s}$, is not a perfect estimate of $Y^0_{t_s}$. This imprecision arises due to two aspects: firstly, $\hat{f}^0$ is not an exact estimate of f^0 thereby generating a prediction error, which we represent as $\mathcal{E}^0_{t_s}(X_{t_s-1}) = f^0(X_{t_s-1}) - \hat{f}^0(X_{t_s-1})$ in the above equation. Secondly, there exist other determinants of the outcome that are orthogonal to the covariates, represented as $u^0_{t_s}$ in the formula. The imprecision emanating from the estimation of f^0, which could differ depending on a firm's characteristics X_{t_s-1}, can be mitigated by exploring various ML techniques and employing the one that delivers the best out-of-sample performance.[4]

In the application we present in this paper, we rely on the "K-fold" cross-validation method (with $K = 5$) to discriminate between the considered ML techniques. We randomly divide the set of exporters observed in $t_s - 2 = 2018$ (considering the exporting success during the same month in $t_s - 1 = 2019$ as the outcome) into 5 equally sized groups and obtain the predictions for the firms belonging to a group by estimating $Y_{2019} = f^0(X_{2018}) + u^0_{2019}$ with different ML models on the firms belonging to the other groups. Then we compute the accuracy of the different models for each month and choose the model with the best average performance across months. Notice that this comparison is entirely based on the pre-pandemic accuracy of the ML models by comparing the predictions $\hat{Y}_{2019}$ with the observed Y_{2019}, not on its merits in predicting the firms' outcomes in 2020. Finally, we obtain the $\hat{Y}^0_{2020}$ by estimating $Y_{2019} = f^0(X_{2018}) + u^0_{2019}$ on the entire set of 2018 units (also in this case month by month) and, as shown in (2.3), applying the estimated function $\hat{f}^0$ to the set of 2019 units. Given that during the first three months of 2020 Colombia was in practice not exposed to COVID-19 (and therefore $Y_{2020} = Y^0_{2020}$), if assumption (2.2) holds, we expect that in those months the accuracy of the predictions $\hat{Y}_{2019}$ obtained in the cross-validation step for 2019 will be very similar to the accuracy of $\hat{Y}^0_{2020}$ for 2020.

[4]In order to simplify the notation, from now on, we will denote $\mathcal{E}^d_{t_s}(X_{t_s-1})$ as $\mathcal{E}^d_{t_s}$.

Following Cerqua and Letta, 2020 and Fabra, Lacuesta, and Souza, 2022, we define as an estimator of the individual-specific shock effect α the simple comparison of the observed outcome under the shock in t_s with the estimated counterfactual outcome. This comparison is represented as:

$$\hat{\hat{\alpha}} = Y_{t_s} - \hat{Y}^0_{t_s}. \tag{2.4}$$

This provides the full distribution of treatment effects.

Starting from Eq. (2.4), by taking the expected value of the individual treatment effect $\hat{\hat{\alpha}}$ for those units with $X_{t_s-1} = x_{t_s-1}$, we can define the following estimator of the conditional average treatment effect (CATE; the average effect for those units with $X_{2019} = x_{2019}$) as

$$\begin{aligned}\mathbf{E}[\hat{\hat{\alpha}}|X_{t_s-1} = x_{t_s-1}] &= \mathbf{E}[(Y_{t_s} - Y^0_{t_s}) - \mathcal{E}^0_{t_s} - u^0_{t_s}|X_{t_s-1} = x_{t_s-1}] = \\ &= \underbrace{\Delta(X_{t_s-1} = x_{t_s-1})}_{\textit{CATE}} - \mathbf{E}[\mathcal{E}^0_{t_s}|X_{t_s-1} = x_{t_s-1}] - \underbrace{\mathbf{E}[u^0_{t_s}|X_{t_s-1} = x_{t_s-1}]}_{\textit{=0 by assumption}},\end{aligned}$$

where,

$$\Delta(X_{t_s-1} = x_{t_s-1}) = \mathbf{E}[Y_{t_s} - Y^0_{t_s}|X_{t_s-1} = x_{t_s-1}]. \tag{2.5}$$

Therefore $\mathbf{E}[\hat{\hat{\alpha}}]$ identifies the unconditional average treatment effect, $\mathbf{E}[\Delta(X_{t_s-1})] = \Delta$, if on average the prediction error is zero: $\mathbf{E}[\mathcal{E}^0_{t_s}] = 0$. The conditional average treatment effect, $\Delta(X_{t_s-1} = x_{t_s-1})$, is identified by $\mathbf{E}[\hat{\hat{\alpha}}|X_{t_s-1} = x_{t_s-1}]$ if on average the prediction error is zero in the relevant sub-sample: $\mathbf{E}[\mathcal{E}^0_{t_s}|X_{t_s-1} = x_{t_s-1}] = 0$.

Now let us decompose the outcome observed in t_s in the presence of the shock, $Y^1_{t_s}$, in a generic model or function $f^1(X^1_{t_s-1})$, which represents the relationship between explanatory variables and the outcome during the shock, and other determinants of the outcome, $u^1_{t_s}$, that are orthogonal to the covariates

$$Y^1_{t_s} = f^1(X^1_{t_s-1}) + u^1_{t_s}, \quad \textit{s.t. } \mathbf{E}[Y^1_{t_s}|X^1_{t_s-1}] = f^1(X^1_{t_s-1}). \tag{2.6}$$

Given that $Y^1_{t_s} = Y_{t_s}$ and $X^1_{t_s-1} = X_{t_s-1}$, then

$$Y_{t_s} = f^1(X_{t_s-1}) + u^1_{t_s}, \quad \textit{s.t. } \mathbf{E}[Y_{t_s}|X_{t_s-1}] = f^1(X_{t_s-1}). \tag{2.7}$$

At this point, we can define an alternative estimator of the individual-specific shock effect α as the comparison of the predicted outcome under the shock in t_s with the estimated counterfactual outcome for a given firm:

$$\hat{\alpha} = \hat{Y}_{t_s} - \hat{Y}^0_{t_s}, \tag{2.8}$$

where $\hat{Y}_{t_s} = \hat{f}^1(X_{t_s-1}) = Y_{t_s} - \mathcal{E}^1_{t_s} - u^1_{t_s}$. We call "Shock Aware Machine" (SAM) the model that we use to predict Y_{t_s} (and the predictions $\hat{Y}_{t_s}$ themselves). The term "Shock Aware" is employed because this model leverages information from the observed shock scenario. Importantly, the predictions from SAM are generated in a metric that allows them to be directly compared with the estimated outcomes (that do not account for the shock), which are produced by the SUM.[5]

In our application, the SAM expresses the outcome in 2020 of exporters operating the foreign market in 2019 as a function of their characteristics in 2019 and the information about governments' shock-related stringency measures all over the world coming from Hale et al., 2021.[6] Similarly to the procedure followed to select the best-performing SUM, we rely on a 5-fold cross-validation strategy to obtain a 2020 prediction for each firm that exported in 2019. We randomly group the 2019 exporters into five equally sized subsets and we predict the 2020 outcomes of the firms contained in one subset by using the information of firms contained in the remaining four subsets. In other words, we train the models on a random 80% of the data and test them on the remaining 20% and we repeat the process five times for each different 20% subset, thus obtaining a 2020 prediction for each 2019 exporter.

Starting from Eq. (2.8), by taking the expected value of the individual treatment effect $\hat{\alpha}$ for those units with $X_{t_s-1} = x_{t_s-1}$, we can define the following alternative estimator of the conditional average treatment effect (for those units with $X_{t_s-1} = x_{t_s-1}$)

$$\begin{aligned}\mathbf{E}[\hat{\alpha}|X_{t_s-1} = x_{t_s-1}] =& \mathbf{E}[(Y_{t_s} - Y^0_{t_s}) - (\mathcal{E}^1_{t_s} - \mathcal{E}^0_{t_s}) - (u^1_{t_s} - u^0_{t_s})|X_{t_s-1} = x_{t_s-1}] \\ =& \underbrace{\Delta(X_{t_s-1} = x_{t_s-1})}_{CATE} - \mathbf{E}[\underbrace{(\mathcal{E}^1_{t_s} - \mathcal{E}^0_{t_s})}_{\Delta\mathcal{E}}|X_{t_s-1} = x_{t_s-1}] - \\ & \mathbf{E}[u^1_{t_s} - u^0_{t_s}|X_{t_s-1} = x_{t_s-1}].\end{aligned} \tag{2.9}$$

[5]Notice that with $\hat{\hat{\alpha}}$, we are comparing a probability (counterfactual) with a binary value (observed outcome), while with $\hat{\alpha}$, we are comparing two estimated probabilities.

[6]We use the firm-level variables "Containment Stringency Index Export" and "Containment Stringency Index Import" which combine the country level information on stringency measures with the exposure of a given firm to the different markets at the export and at the import side in $t_s - 1$, as explained in more detail in subsection 2.3.1. We do not introduce these variables explicitly as an argument of $f^1()$ to simplify notation.

Therefore, $\mathbf{E}[\hat{\alpha}]$ identifies the unconditional average treatment effect, $\mathbf{E}[\Delta(X_{t_s-1})] = \Delta$, if, on average, the difference in prediction errors is zero: $\mathbf{E}[\Delta\mathcal{E}] = 0$. The conditional average treatment effect, $\Delta(X_{t_s-1} = x_{t_s-1})$, is identified by $\mathbf{E}[\hat{\alpha}|X_{t_s-1} = x_{t_s-1}]$ if on average the difference in prediction errors is zero in the relevant sub-sample: $\mathbf{E}[\Delta\mathcal{E}|X_{t_s-1} = x_{t_s-1}] = 0$.

Given the definitions of SUM and SAM, to simplify the reasoning in the following, we refer to Eqs. (2.4) and (2.8), respectively as

$$\hat{\hat{\alpha}} = Y - \hat{Y}_{SUM} = Y - SUM. \tag{2.10}$$

$$\hat{\alpha} = \hat{Y}_{SAM} - \hat{Y}_{SUM} = SAM - SUM. \tag{2.11}$$

The assumptions behind these identification results are not directly testable as they are expressed in terms of the expected values of the prediction error $\mathcal{E}^0_{t_s}$ that is a function of the unobservable counterfactual $Y^0_{t_s}$. Table 4 distinguishes the five different scenarios concerning the values of $\mathcal{E}^0_{t_s}$ and $\mathcal{E}^1_{t_s}$ that are relevant in determining whether applying the statistic $\mathbf{T}$ to $Y - SUM$ and $SAM - SUM$ is able to recover the corresponding treatment effect estimand (e.g., whether averaging the estimated individual treatment effects would recover the average treatment effect).

Table 4: Identification of generic functions of the individual treatment effects, $\mathbf{T}$, according to the corresponding value taken by the prediction errors.

	$\mathbf{T}(SAM - SUM)$	$\mathbf{T}(Y - SUM)$
$\mathbf{T}[\mathcal{E}^1_{t_s}] \neq 0$ and $\mathbf{T}[\mathcal{E}^0_{t_s}] = 0$	×	✓
$\mathbf{T}[\mathcal{E}^1_{t_s}] = \mathbf{T}[\mathcal{E}^0_{t_s}] = 0$	✓	✓
$\mathbf{T}[\mathcal{E}^1_{t_s}] = 0$ and $\mathbf{T}[\mathcal{E}^0_{t_s}] \neq 0$	×	×
$\mathbf{T}[\mathcal{E}^1_{t_s}] = \mathbf{T}[\mathcal{E}^0_{t_s}] \neq 0$	✓	×
$\mathbf{T}[\mathcal{E}^1_{t_s}] \neq \mathbf{T}[\mathcal{E}^0_{t_s}] \neq 0$	×	×

The estimators based on $Y -$SUM identify the population parameters when $\mathbf{T}[\mathcal{E}^0_{2020}] = 0$. The estimators based on $SAM - SUM$ are unbiased whenever $\mathbf{T}[\mathcal{E}^1_{2020}] = \mathbf{T}[\mathcal{E}^0_{2020}]$. Under the assumption that the strength of the COVID-19 effect on export propensity was at most very limited during the first quarter of 2020, we will use the out-of-sample prediction errors for the first quarter of 2020 as a proxy for the unobservable behavior of $\mathcal{E}^0_{2020}$ in the following months. Moreover, as explained in detail in section 2.4.2, the distribution of the estimated treatment effects during the first quarter will be used to check the credibility of the above assumptions for the set of all 2019 exporters and for different subsets of 2019 exporters defined according to their characteristics X_{2019} or to their position in the distribution of such effects.

As a final step, we perform the heterogeneity analysis by adapting the Sorted Partial Effect (SPE) method introduced in Chernozhukov, Fernández-Val, and Luo, 2018. Formally, the SPEs are defined as percentiles of the Treatment Effects (TE) and can supply a more detailed summary of the distribution of TE than the Average Treatment Effects (ATE), commonly employed in econometric analysis. The SPEs are defined as

$$\alpha^*(u) = u^{th} - \text{ percentile of } \alpha. \tag{2.12}$$

In our setting, $\alpha^*(u)$ is a function of X_{t_s-1} defined over its distribution in the population of $t_s - 1$ exporters.

The SPEs are used to do a classification analysis (CA) that allocates the $t_s - 1$ exporters into two groups, the most and the least affected by the shock, according to whether their α are lower than $\alpha^*(25)$ or greater than $\alpha^*(75)$, respectively. Notice that, since the shock effect is negative, we have defined as the most (negative) affected units those whose α lie in the left tail of the sorted distribution of treatment effects. Finally, to study the determinants of treatment effect heterogeneity, we focus on the difference in means ($CADiff$) of the X_{t_s-1} across the most and least affected groups. In the estimation, we use sample analogues of $\alpha^*(u)$ and $CADiff$. We calculate standard errors of $\alpha^*(u)$ and $CADiff$ by bootstrapping the entire estimation process, starting from the initial α estimation step.

The application of the SPE technique presents several advantages in our setting. First, the estimated $\alpha^*(u)s$ provide a summary of the distribution of the estimated treatment effects and, therefore, of treatment effect heterogeneity. Second, it can be used to identify a subgroup of the population that is more affected by the treatment and to study how the heterogeneity of the treatment effect depends on observables without imposing (additional) functional form assumptions. Third, it provides p-value adjustments to account for joint testing of all the covariates considered in the $CADiff$ final step, which is aimed at detecting if observables are associated with treatment effect heterogeneity. In other words, the main idea is to test the null hypothesis of no difference between the value of the covariates in the most and the least affected groups because we conduct simultaneous inference on multiple variables.[7]

Our methodology for estimating individual treatment effects and conducting heterogeneity analysis closely aligns with the general machine learning framework proposed by Chernozhukov et al., 2020. However, our approach has been specifically adapted to address scenarios where a contemporaneous control group is absent—i.e., where identifying ex-ante firms unaffected by the shock is challenging. This adaptation, as further detailed in Section 2.2.1, allows for robust inference despite the lack of a clear control group, making it particularly suited for the unique characteristics of our empirical setting.

2.2.1 Methodological comparison with Chernozhukov et al., 2020

We demonstrate that our empirical strategy is grounded in the same foundational principles as those outlined by Chernozhukov et al., 2020, yet it is applied to a distinct empirical context. To aid in clarifying our approach, we refer to Figure 11, which offers a simplified illustration of our specific empirical setting.

Chernozhukov et al., 2020 deal with an experimental empirical setting in which one can easily separate a treated group from a control group. In order to study the heterogeneity of the average treatment effect, the first step of Chernozhukov et al., 2020 is to split randomly the sample under analysis in an auxiliary (A) and a main sample (M) of approximately the same size. Then, they employ ML techniques to learn in A the function approximating the potential outcomes in the treatment and non-treatment scenarios, while M is used to make inferences on the key features of treatment effect heterogeneity. In other words, they estimate the function describing the outcome in case of treatment (no treatment) on the subset of treated (non-treated) firms contained in A. These

[7] Starting from B bootstrap replications of all the estimation steps (including the prediction stage), we calculate the $CADiff$ B times. To determine the significance of the $CADiff$ we perform a two-tailed test. The p-values are constructed as follows:

$$2 \cdot min\{Pr(S \geq t|H_0), Pr(S \leq t|H_0)\}$$

being t the observed t test statistic, $t = \frac{\widetilde{CADiff_{original}}}{\tilde{\sigma}}$, drawn from the unknown distribution S. $\tilde{\sigma}$ represents the standard deviation of the bootstrapped $CADiff$. To adjust the p-values and obtain the joint p-values taking into account that we are testing hypotheses jointly on many covariates, we reproduce the "single-step" method employed by Chernozhukov, Fernández-Val, and Luo, 2018 to control for the family-wise error rate.

two estimated functions are used to impute the two potential outcomes for each firm contained in the M sample (the difference represents the estimated individual treatment effects) and study the treatment effect heterogeneity estimated for these firms by using, inter alia, the Sorted Effects method (Chernozhukov, Fernández-Val, and Luo, 2018). This procedure is designed in this way to avoid overfitting (i.e., doing learning and prediction using the same sample), and, starting from the random splitting, it is repeated many times in order to obtain many distributions of estimated treatment effects to which the Sorted Effects method is applied.

Figure 11: Empirical Setting - SUM vs. SAM vs. Chernozhukov et al., 2020

Our Setting			Chernozhukov (2020) Setting	
SUM		SAM	$A-M$ splitting	
$(X_{2018}, Y_{2019})_1$	$(X_{2019}, Y_{2020})_{11}$	$(X_{2019}, Y_{2020})_{11}$	$(X_{2019}, Y_{2020})_{11}$	
$(X_{2018}, Y_{2019})_2$	$(X_{2019}, Y_{2020})_{12}$	$(X_{2019}, Y_{2020})_{12}$	$(X_{2019}, Y_{2020})_{12}$	
$(X_{2018}, Y_{2019})_3$	$(X_{2019}, Y_{2020})_{13}$	$(X_{2019}, Y_{2020})_{13}$	$(X_{2019}, Y_{2020})_{13}$	A
$(X_{2018}, Y_{2019})_4$	$(X_{2019}, Y_{2020})_{14}$	$(X_{2019}, Y_{2020})_{14}$	$(X_{2019}, Y_{2020})_{14}$	
$(X_{2018}, Y_{2019})_5$	$(X_{2019}, Y_{2020})_{15}$	$(X_{2019}, Y_{2020})_{15}$	$(X_{2019}, Y_{2020})_{15}$	
$(X_{2018}, Y_{2019})_6$	$(X_{2019}, Y_{2020})_{16}$	$(X_{2019}, Y_{2020})_{16}$	$(X_{2019}, Y_{2020})_{16}$	
$(X_{2018}, Y_{2019})_7$	$(X_{2019}, Y_{2020})_{17}$	$(X_{2019}, Y_{2020})_{17}$	$(X_{2019}, Y_{2020})_{17}$	
$(X_{2018}, Y_{2019})_8$	$(X_{2019}, Y_{2020})_{18}$	$(X_{2019}, Y_{2020})_{18}$	$(X_{2019}, Y_{2020})_{18}$	M
$(X_{2018}, Y_{2019})_9$	$(X_{2019}, Y_{2020})_{19}$	$(X_{2019}, Y_{2020})_{19}$	$(X_{2019}, Y_{2020})_{19}$	
$(X_{2018}, Y_{2019})_{10}$	$(X_{2019}, Y_{2020})_{20}$	$(X_{2019}, Y_{2020})_{20}$	$(X_{2019}, Y_{2020})_{20}$	

Note: A simplified representation of our empirical setting in which we compare the methods used in the present paper to those described in Chernozhukov et al., 2020.

As an example, in Figure 11, we represent 20 exporting firms observed in 2018 or in 2019. In the context of our setting, the strategy of (Chernozhukov, Fernández-Val, and Luo, 2018) would imply that the 2019 sample, for which we are interested in estimating the average treatment effect, should be divided in two, as shown in the last column of Figure 11. However, in the COVID-19 scenario, one cannot easily separate treated and control units because COVID-19 imposes a, at least indirectly, treatment over all units, hence preventing the possibility of discerning between treated and controls.[8]

[8]Furthermore, even if we assume that during the first three months of the year there was

Moreover, with respect to Chernozhukov et al., 2020, in our empirical setting, we do not have the necessity to predict the outcome of controls in the case of "no treatment" because we are not interested in estimating the COVID-19 effect on 2018 exporters. Therefore, we do not have to split the controls observed in 2018 in two halves to avoid overfitting and this enables us to reconstruct a counterfactual outcome of no treatment for each 2019 exporter without incurring in overfitting problems. Therefore, in this paper for the SUM we use as an auxiliary sample all the Colombian exporters observed in 2018 (A) and as the main sample (M) all the Colombian exporters observed in 2019. For the SAM, we perform instead a K-Fold splitting in which, iteratively we select 80% of the firms in 2019-2020 as being part of A and the remaining 20% as being part of M. This is shown in the column "Our Setting (SAM)" of Figure 11, where different A (and, accordingly different M) groups are selected according to the different colors of the dashed circles. In this way we avoid overfitting problems and, at the same time, we exploit all the available data by being able to compare the predicted probabilities to export in the COVID-19 with those in the non-COVID-19 scenario for all the observed 2019 exporters.

2.3 Data

This study focuses on the social and economic disruption caused by the COVID-19 pandemic and its effect on Colombian exporters. This global health crisis served as a notable example of a large-scale economic shock that profoundly impacted global trade, with the dynamics of exporters in Colombia being significantly affected. Applying our machine learning strategy to data collected from Colombian exporting firms during this period can yield substantial insights into how such entities adapt and persist in the face of such extensive disruption. This provides an understanding of market resilience and firm survival dynamics in the context of global trade shocks. By grounding our research in a tangible case

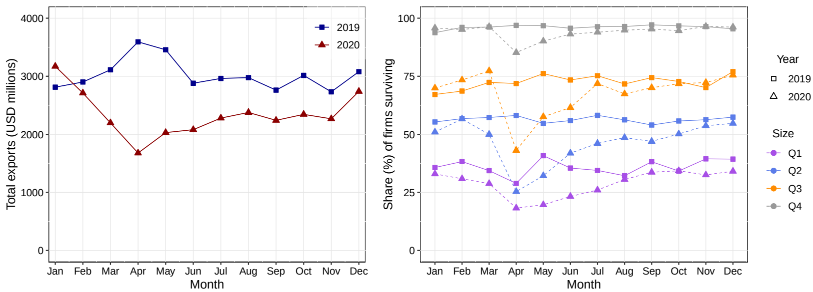
no COVID-19 effect going on, and therefore we categorize as non-treated (treated) firms operating in those months (in the other remaining months), and we use the non treated firms in the auxiliary sample to learn, it would be problematic to use the learning outcome in case of no treatment during the first three months to predict the outcome in case of no treatment for the treated firms that are those in the last 8 months because of the strong seasonality effects we have. So the outcome during the first three months in case of no treatment would be very different from the outcome of the last months in case of no treatment just because of seasonality effects.

study, we maintain its relevance to the specific scenario while preserving its potential for broader applications.

We use monthly export transaction data reported at the Colombian Customs Office (Dirección de Impuestos y Aduanas Nacionales, DIAN) for 2018, 2019, and 2020. For each transaction, we consider the exporter ID as the firm identifier; the date; a 10-digit Harmonized System code (HS) characterizing the product; the product origin within Colombia (department level); the means of transportation of the shipment; the country of destination; and, the free on board value of the export transaction in US dollars. This data set also contains information about the value and origin country from which a given exporter imports. We remove all transactions related to re-exports of products elaborated in other countries. As a result, we ended up with 386,132 customs reports in 2018 (7741 firms), 402,140 in 2019 (7831 firms), and 365,626 in 2020 (7518 firms).

The left panel in Figure 12 shows the evolution of total monthly exports during 2019 and 2020. The total monthly value of exports in 2020 is significantly lower than the one observed for the corresponding month in 2019, except for January and February. The lockdown measures implemented to contain the COVID-19 outbreak in Colombia and abroad had a severe impact between April and June–the value in April 2020 is half of the one observed in April 2019 (47%).

Figure 12: Evolution of total exports and proportion of surviving exporters.



Note: The evolution of total exports (left) and the proportion of surviving exporting firms in year t among those exporting in year $t-1$ within size classes defined at $t-1$ (right). Firm size classes are based on the quartiles of the firms' exports (in ln) distribution in a given month.

In a typical month, large firms get a lion's share of the total exports. A regular pattern in looking at customs data is that more prominent exporters trade for many months and ship more frequently than smaller firms, which make only a few shipments. The right panel in Figure 12 shows the proportion of surviving exporting firms in year t among those exporting in year $t-1$, by size classes defined at $t-1$. Comparing the figures for 2020 with those for 2019, it seems that the COVID-19 outbreak affected all firms regardless of their size. However, the effect looks proportionally stronger for small firms (Q1 and Q2 of the distribution). In contrast, larger firms are less affected and recover faster than the survival rates observed in 2019.

Appendix B.3 shows the growth patterns of the number of exporters and export volumes between 2019 and 2020 (and, as a comparison, between 2018 and 2019) segmented by country of destination and product sector, offering further insights into the heterogeneous impacts of the COVID-19 pandemic on Colombian exports.

2.3.1 Control Variables

We classify products at the six-digit level of the HS code. We consider different features of exporters according to their monthly exports: the total export (and import) value, the number of products (NP), the number of export destinations (ND), the number of import origin countries (NO), the Herfindahl-Hirschman indexes at the product level (HH_p) and the destination level (HH_d), and a set of dummies for the destinations and origin countries and continents. We create a set of dummies according to the Colombian department from which the product comes, a set of dummies for the means of transportation used, and a set of dummies classifying the product HS-chapter and HS-section. Moreover, we build two sets of dummy variables indicating whether a firm has experience exporting in specific destinations and product sectors. We also account for the accumulated exporting (importing) experience by summing up the total value exported (imported) during the last twelve months. Furthermore, we create four *size* dummies classifying firms according to the quartiles of the firm-level distribution of the total monthly log-value of exports.

To measure the COVID-19 demand and supply shock, we use the information on government contention measures coming from Hale et al., 2021, which consists of four indexes (ranging from 0 to 100) representing the strength of the measures taken by countries to contain the COVID-19 outbreak. The authors provide an economic index summarizing economic policies (E), a health index summarizing health policies (H), a government index describing the strictness of 'lockdown style' policies (G), and an overall government response index called stringency index (S). The value of these indexes ranges from 0 to 100.[9] We build two variables at the firm level for each of the four indexes, one at the export and one at the import side, by taking a weighted average of the country-level scores according to the proportion of the total monthly value of exports (imports) that a firm ships (source) in each country in 2019. We call these firm-level indexes for a firm i "Containment Index$_{i,j,z}$", with $j = \{E, H, G, S\}$ and $z = \{\text{Imp}, \text{Exp}\}$.[10]

Our final data set is composed of 1,975 covariates. They are presented in detail in Table 17 of Appendix C.1.

2.4 Results

2.4.1 Selection of the machine learning algorithm

We evaluate and contrast the outcomes of several ML techniques against a benchmark logistic regression, aiming to identify the model with superior prediction performance. The out-of-sample predictive efficacy of our empirical models is crucial, given our goal to reconstruct an unobserved counterfactual. The complexity of this task arises from its high dimensionality and complex interdependencies between firms and products from various sectors and export destinations. While an approach focusing on in-sample prediction accuracy might overfit, ML techniques optimally balance the bias-variance trade-off for out-of-sample predictions.[11]

We examine four distinct models: Logit, Logit-LASSO, Logit-Ridge, and Random Forest (RF). The traditional choice for binary dependent variables, Logit, serves as our baseline. Even though literature often shows ML techniques outperforming traditional models with numerous predictors, we have included Logit results for comparison. The main idea of Logit-LASSO is to mitigate overfitting by introducing a penalty term in the Logit log-likelihood function that forces the parameters associated with the less relevant predictors to be exactly zero. On the other

[9] These indexes are released daily. We average this information at the monthly level.

[10] The value of the Containment Stringency Index Import for firms that are not importing corresponds to the value of the Containment Stringency Index for Colombia (as firms are sourcing all their inputs domestically).

[11] Hyperparameter tuning through cross-validation or other theory-driven methods is often critical in order to avoid overfitting.

hand, Logit-Ridge reduces the coefficients of less significant predictors without eliminating any of them, proving especially useful when many variables play an important role. The main idea behind Random Forest is the wisdom of crowds because it combines the predictions of many uncorrelated models (the trees) obtained by randomly re-sampling observations and explanatory variables.[12] For Logit, Logit-Ridge, and Logit-LASSO models we include interactions between the *size* of the company and some of the main product characteristics, *industry*, *sector*, *means of transportation* as well as with *destination country* dummies. Notice that Random Forest uses the variables sequentially and, therefore, with a large enough number of trees, it is not necessary to explicitly introduce interactions as explanatory variables, i.e., the model automatically takes into account the interactions that are useful to accurately predict the outcome.[13] The prediction analysis is repeated for all months between January-December 2020.

Table 5 shows the accuracy of the model's predictions through two widely-used classification performance metrics: Area Under the Receiver Operating Curve (AUC) and Root-Mean-Square Error (RMSE). The AUC achieves a value of 0.5 for random predictions and 1 when outcomes are classified without error. Meanwhile, RMSE's best score is 0, indicating optimal accuracy with no fixed upper bound.

The table's upper part displays the accuracy of predictions for the probability of exporting in 2019 based on 2018 exporter data, serving as an out-of-sample performance benchmark in a pre-COVID-19 context using cross-validation. Here, the Logit-LASSO and RF models arise as top performers. The table's middle section also shows the accuracy of models estimated using the exporters' characteristics in 2018 to explain their observed outcomes in 2019; however, these models are now tested using the set of exporters of 2019 and their observed outcomes in 2020. If the

[12]Note that it is important to optimize (tune) the hyper-parameters of Logit-Ridge and Random Forest for an accurate predicting exercise. The hyperparameter to optimize in Logit-Ridge and Logit-LASSO is λ, which controls the impact of the penalty or shrinkage on parameters (when $\lambda = 0$ we are in a Logit scenario when $lambda$ increases the penalty impact grows). We find the optimal hyper-parameter for Logit-Ridge and Logit-LASSO by choosing the λ that minimizes the mean cross-validated error. One of the main RF parameters is the number of random trees used. Because of the RF design, it is very difficult to have over-fitted predictions when using this model. Therefore, we set the number of trees to a large enough number (500). We find this number is large enough after repeating the same Random Forest model with a different number of random trees. Even with less than 500 random trees, the error rate of the model remains unchanged.

[13]For more information about all the features included to build the SUM and SAM see Table 17 in Appendix C.1.

functions f_t^0 representing the relationship between explanatory variables and the outcome in the absence of the pandemic are sufficiently similar for the pre-pandemic year and 2020 (f_{2019}^0 and f_{2020}^0, respectively), we expect that the accuracy of $\hat{f}_{2019}^0$ in the first three months of 2020 (when arguably no relevant COVID-19 effect is in place in Colombia) to be similar to that one which is observed during the same months of 2019. Indeed, during January, February, and March, the accuracy of Logit-LASSO and RF remains unchanged, as expected, compared to the accuracy obtained in the upper part of the table. However, a decline in accuracy appears post-April in the middle part of the tables, demonstrating the challenges of a model not trained on COVID-19 data, predicting in a COVID-19-impacted environment.

Models in the bottom part of Table 5 are trained and tested with the universe of exporters in 2019 and their observed outcomes in 2020. Using these models, we construct the SAM predictions. The accuracy of the predictions is very similar to the one obtained with the SUM for 2019 and for the first three months of 2020. Our analysis is crucial to reach accurate predictions because the unbiasedness of our treatment effects estimators depends on the quality of the (counterfactual) prediction accuracy. Both the SUM and the SAM show acceptable levels of accuracy when predictions are made with Logit-LASSO and Random Forest.

2.4.2 Evaluation of the COVID-19 effect

Both Logit-LASSO and Random Forest reach high accuracy levels in the export status prediction. As explained in section 2.2, the predicted probabilities are used to estimate the average monthly effect of the COVID-19 shock as the monthly average of $\hat{\alpha}$ (the difference between the firm-level probabilities of success predicted by the SUM and the SAM.). They are presented in Figure 13.

Given the presumption that firms suffered a negligible COVID-19 shock impact during the initial three months of 2020, the treatment effect estimates for this period can be viewed as a placebo test, reminiscent of the in-time placebo test routinely employed in Synthetic Control Methods (Abadie, Diamond, and Hainmueller, 2015). Detecting a significant COVID-19 effect in the months preceding the actual economic shock would suggest that our model mechanically estimates a COVID-19 effect even in the absence of the stated shock. We conduct these placebo studies also conditioning on exogenous firms' characteristics observed in 2019 by

Table 5: Goodness of Fit for SUM and SAM in 2018/19 and 2019/20

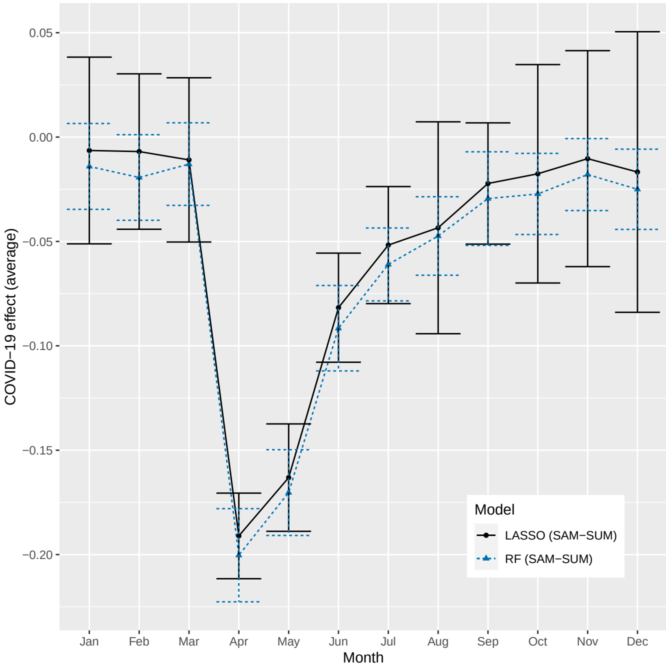
	AUC				RMSE			
	Logit-LASSO	Logit-Ridge	Random Forest	Logit	Logit-LASSO	Logit-Ridge	Random Forest	Logit
Goodness of Fit for SUM in 2018/19								
Jan	0.73	0.53	0.73	0.59	0.40	0.45	0.41	0.64
Feb	0.70	0.50	0.71	0.58	0.41	0.45	0.41	0.64
Mar	0.70	0.56	0.71	0.57	0.41	0.44	0.41	0.65
Apr	0.73	0.59	0.73	0.60	0.40	0.43	0.40	0.63
May	0.72	0.52	0.71	0.59	0.40	0.44	0.41	0.64
Jun	0.71	0.50	0.72	0.59	0.40	0.45	0.41	0.64
Jul	0.73	0.50	0.73	0.55	0.40	0.45	0.40	0.66
Aug	0.70	0.51	0.72	0.58	0.41	0.45	0.40	0.64
Sep	0.72	0.50	0.71	0.58	0.41	0.45	0.40	0.64
Oct	0.73	0.58	0.74	0.58	0.40	0.44	0.41	0.64
Nov	0.71	0.51	0.72	0.57	0.41	0.45	0.41	0.64
Dec	0.70	0.50	0.71	0.58	0.41	0.45	0.41	0.64
Goodness of Fit for SUM in 2019/20								
Jan	0.72	0.53	0.72	0.49	0.41	0.45	0.41	0.75
Feb	0.69	0.50	0.69	0.56	0.41	0.45	0.42	0.64
Mar	0.72	0.54	0.73	0.59	0.40	0.44	0.41	0.63
Apr	0.67	0.56	0.66	0.51	0.48	0.50	0.49	0.70
May	0.69	0.51	0.69	0.60	0.46	0.48	0.46	0.63
Jun	0.68	0.50	0.68	0.59	0.43	0.47	0.44	0.63
Jul	0.70	0.50	0.69	0.59	0.42	0.46	0.43	0.63
Aug	0.68	0.51	0.69	0.58	0.42	0.45	0.43	0.63
Sep	0.69	0.50	0.70	0.59	0.42	0.45	0.42	0.63
Oct	0.71	0.59	0.70	0.60	0.42	0.45	0.43	0.63
Nov	0.71	0.51	0.71	0.59	0.41	0.45	0.41	0.63
Dec	0.69	0.50	0.69	0.58	0.42	0.46	0.42	0.63
Goodness of Fit for SAM in 2019/20								
Jan	0.73	0.58	0.74	0.50	0.41	0.45	0.41	0.71
Feb	0.70	0.50	0.70	0.49	0.41	0.46	0.42	0.70
Mar	0.73	0.50	0.73	0.50	0.40	0.46	0.40	0.71
Apr	0.74	0.66	0.73	0.52	0.42	0.47	0.42	0.69
May	0.76	0.74	0.77	0.50	0.41	0.46	0.41	0.71
Jun	0.73	0.69	0.73	0.48	0.42	0.46	0.42	0.72
Jul	0.73	0.63	0.72	0.51	0.41	0.45	0.42	0.69
Aug	0.72	0.50	0.72	0.53	0.41	0.46	0.42	0.69
Sep	0.71	0.50	0.70	0.55	0.42	0.47	0.42	0.67
Oct	0.72	0.50	0.71	0.52	0.42	0.46	0.42	0.70
Nov	0.72	0.52	0.72	0.49	0.41	0.45	0.41	0.71
Dec	0.71	0.51	0.70	0.51	0.41	0.45	0.42	0.70

estimating COVID-19 effects for selected subsamples of firms according to such characteristics. We interpret these additional placebo studies as a robustness check on our results on treatment effect heterogeneity.

As shown in Figure 13, the probabilities obtained from the SUM and the SAM are almost identical on average for January, February, and March. This result is reassuring since only from March 25, 2020, the Colombian government implemented a complete and mandatory lockdown. More in general, we can conclude that our identification strategy is not mechanically recovering COVID-19 effects for a period with low incidence in Colombia and in the rest of the world. We find that the peak of the COVID-19 effect is in April 2020, when we estimate an average difference between the predicted probabilities of exporting of nearly 20 percentage points. In the following months, the estimated average effect declines with time.

The results indicate that both Logit-LASSO and RF models yield comparable performances. Given their good performance and considering that Logit models are frequently used in similar contexts, we opt for

Figure 13: Average Treatment Effect (ATE)



Note: Average Individual Treatment Effect, by months, comparing Logit-LASSO and RF. Standard errors were obtained with 100 bootstrap replications. Confidence intervals for a 5% significance level.

Logit-LASSO. It aligns with the conventional approaches and offers greater interpretability as an extension of the traditional model.[14]

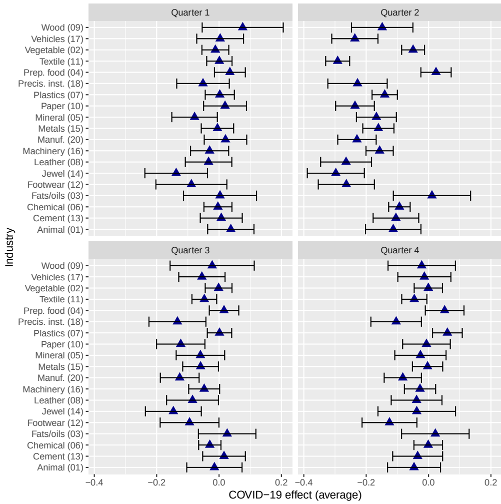
Figure 14 shows evidence of substantial variations in the quarterly estimated average individual treatment effect by industry. On the one hand, during the first, third, and fourth quarters of 2020, there is no evidence supporting the existence of sectoral heterogeneity in the COVID-19 effect, and the COVID-19 shock is economically and statistically insignificant. Therefore, concentrating on the results for the first quarter, we are able to reject the existence of an effect even within sectors.[15]

[14]Non-reported results using RF are equivalent and available upon request.

[15]We have conducted other similar placebo studies conditioning on other variables (e.g.,

On the other hand, during the second quarter of 2020, Colombian exporters belonging to almost every industry are found to significantly reduce their probability of surviving in the international markets. This decline is particularly pronounced in industries such as Textiles, Footwear, and Jewelry. However, industries like Food Preparations and Vegetables saw minimal changes in their survival probabilities due to the COVID-19 shock.

Figure 14: ATE by industry



Note: The quarterly mean difference in the predicted probability of success (SAM vs. SUM) by industry, using the Logit-LASSO predictions. Standard errors were obtained with 100 bootstrap replications. Confidence intervals for a 5% significance level.

2.4.3 Heterogeneity of the COVID-19 effect on Colombian exporters

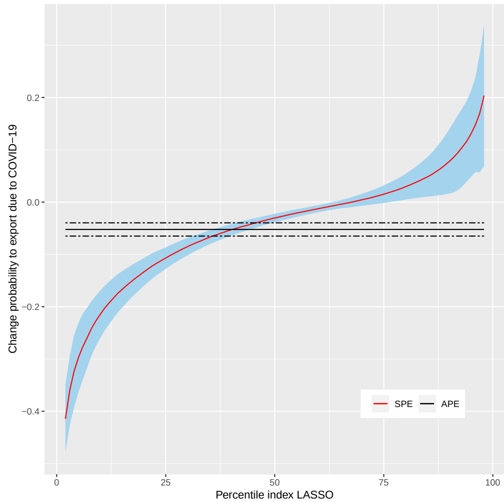
In this section, we investigate the determinant of possible treatment effect heterogeneity. Figures 15 and 16 show the estimated Sorted Partial Effects (SPE) and Average Partial Effects (APE), which are obtained as

---

the main destination of exports, the main origin of imports,...) and in all the considered subsamples we do not estimate any significant effect of COVID-19.

explained in section 2.2 by month and aggregating all the months, respectively. The two figures also report the 95% confidence intervals with blue bands for SPE and black dashed lines for APE.

Figure 15: Annual Sorted Partial Effects (SPE) by percentiles (SAM-SUM)



Note: Annual Sorted Partial Effects (SPE) and Average Partial Effects (APE) of COVID-19 on Colombian firm export's status. The treatment effect is calculated as a difference between SAM and SUM predictions. Standard errors were obtained with 100 bootstrap replications. Confidence intervals for a 5% significance level.

Significant treatment effect heterogeneity is observed for April and May, with June showing a milder effect. The statistically significant (negative) estimated values of $\alpha^*(u)$ are primarily confined to the distribution's left tail. However, from July onwards, the confidence intervals of the SPEs overlap with those of the APEs, indicating an absence of treatment effect heterogeneity. Interestingly, in the pre-pandemic months, the SPEs closely aligned with the APEs (estimated to be zero). This demonstrates that individual placebo treatment effects are not statistically significant throughout the distribution, not just on average, reinforcing the robustness of our methodology across the entire distribution of treatment effects.

To identify the determinants of treatment effect heterogeneity, we examine the difference in means ($CADiff$) of firm characteristics between the most and least affected groups in Table 6. These groups are defined by whether their estimated α is lower than $\alpha^*(25th)$ or greater than $\alpha^*(75th)$, respectively. Therefore, we compute the raw difference in the means of the covariates between the most and the least affected firms by regressing the variables of interest on a constant and a dummy $q = \mathbf{1}_{\{\alpha \leq \alpha^*(25th)\}}$ for the observations for which $\alpha \leq \alpha^*(25th)$ or $\alpha \geq \alpha^*(75th)$. Then, we also provide the difference in adjusted means once we have controlled for the firm sector and month of the year. Controlling for sector and month allows us to perform a *ceteris paribus* analysis, i.e., to dig into the effects of COVID-19 within specific sectors and specific months.

Table 6 is divided into 3 columns according to the control variables included in the regressions: in the first column, we show the unconditional average difference in the firms' characteristics between the most and least affected firms; in the second column, we control for the firm sector; and, in the third column, we control for firm sector and month of observation. The firm characteristics that we consider to explore the sources of COVID-19 treatment effect heterogeneity among Colombian exporters are observed in 2019 (the year before receiving the treatment). First, we check whether the estimated individual treatment effect (TE) differs between the firms contained in the two groups by using the TE as the dependent variable. We then move to firm-sector specific characteristics. In particular, the first set of firm characteristics that we use as dependent variables are dummies indicating the industry where the exporters operate.[16] We also investigate the *CADiff* for the means of transportation and the months when firms operate. Moreover, to account for the role of diversification patterns, we also consider as dependent variables the number of export destinations (ND), import origins (NO), and products (NP) exported. The weighted Containment Stringency Index that exporters face when exporting (importing) allows us to study to what extent treatment effect heterogeneity depends on these measures of direct exposure (to COVID-19 through their activities on international markets). A traditional continuous-DID strategy would have

[16]We aggregate the 22 industries defined in the main analysis as follows. "Agriculture" contains Animals (01), Vegetables (02), Fats/oils (03), and Prepared Foodstuffs (04). "Chemicals" includes Chemical (06), and Plastics (07). "Manufacturing" contains Machinery (16), Vehicles (17), and Manufactured (20). "Metals" aggregates Mineral (05), Cement (13), Jewelries (14), and Metals (15). "Special" includes Precision Instruments (18), Arms (19), Art (21), and Special (22). "Textile" contains Leather (08), Textile (11), and Footwear (12). Finally, "Wood" aggregates Wood (09), and Paper (10). See Table 16 for the complete industry names.

used these exposure variables as treatment variables, assuming that any COVID effect would emanate through them. Finally, including the total value exported (imported) by firms –expressed in logarithm– among the variables for which the *CADiff* is computed highlights the difference in the quantities sold (purchased) by most and least affected companies. A discussion of the main findings follows.

Considering the estimated individual treatment effects (TE) as a dependent variable, we find a negative and significant difference between most and least affected firms independently of the set of controls employed. These results show that the most affected exporters–i.e., those located in the first SPE quartile distribution–experienced a decrease in the probabilities of exporting between 27.9 and 31.3 percentage points lower than the one experienced by the least affected firms–i.e., those located in the last SPE quartile.

We found significant differences among firms when examining how different aggregate sectors are affected. For instance, we detect that the share of textile firms among the most affected 2019 exporters is 16 percentage points higher with respect to the one estimated for the group of the least affected firms. Likewise, there is a difference of 2.9 percentage points in the presence of wood exporters between the most and least affected groups.

We also detect the existence of treatment effect heterogeneity associated with the means of transportation used by exporters in 2019. On the one hand, there are 16.8 to 20.4 percentage points more exporters using air transportation among the most affected than among the least affected firms. However, there are 19.2 to 23.6 percentage points fewer Colombian exporters using the sea for shipping among the most impacted firms compared to the least affected ones (Nitsch, 2022).

Looking at the treatment effect heterogeneity associated with months, the first pattern we notice is that only the months from April to August have a positive estimated parameter. However, only April and May estimated differences are statistically significant. There are 18.6 to 19.5 percentage points (17.7 to 18.2) more firms in April (May) among the most affected than among the least affected firms. From September to November, the coefficients become negative and significant, indicating the beginning stages of recovery.

To evaluate how ex-ante exporter diversification affects the COVID-19 effect, we explore the estimated parameters associated with ND, NO, and NP. We want to investigate whether Colombian exporters' supply chain diversification and export destination diversification help mitigate the COVID-shock. We do not find compelling evidence that ex-ante diversification helps to face a shock of this kind, as we can evince from the estimated parameters associated with ND, NP, and, in the first column, to NO. Following the reasoning of Lafrogne-Joussier, Martin, and Mejean, 2022, which exploits the COVID-19 crisis to study the export consequences of a country-specific supply-side shock by concentrating on the differential import exposure of French firms to the Chinese early lockdown, one possible explanation is that firms cannot substitute away the partner (or the product) under COVID-19. Another possible explanation, which they offer, is that exporters that do not diversify ex-ante can benefit from some form of ex-post diversification. However, when they restrict the analysis to homogeneous inputs, Lafrogne-Joussier, Martin, and Mejean, 2022 find weak evidence of a larger COVID-19 effect for firms with non-diversified inputs. They restrict the sample to homogeneous inputs because they want to analyze the COVID-19 effect among inputs expected to be substituted. Similarly, once we control for the sector and, therefore, inter alia, for the fact that some sector has relatively more diversification potential, the negative estimated difference turns statistically significant. Indeed, within sectors, the most affected Colombian exporters tend to import from 1.98 fewer countries in 2019 than the least affected firms. The economic size of this estimate is large as approximately 60 per cent of Colombian exporters are not integrated into global value chains (they do not import), and the mean of NO is approximately 4.16 origins.

The $CADiff$ estimated when using the Export (Import) Containment Stringency Index as dependent variables provides insightful hints on the difficulties of Colombian firms in exporting (importing) to (from) countries adopting severe stringency measures. In particular, the most affected Colombian exporters face, on average, a higher Export (Import) Containment Stringency Index than those faced by least affected firms by 7.18 to 19.51 (7.25 to 20.80) points, depending on the column in the table.[17]

Finally, the least affected firms exported (imported) 156.7% to 176.83% (614.7% to 1467.3%) more value in 2019 than the most affected firms. As expected, Colombian exporters trading in larger volumes (in value) are more resilient under a COVID-19 scenario. As with diversification, the comparison of the export and import side reinforces the idea that having more experience in sourcing inputs from abroad decreases the strength

[17] Remember that the Index ranges from 0 to 100.

of the shock.

2.4.4 Estimations based on $Y - SUM$

In this paragraph, following Fabra, Lacuesta, and Souza, 2022 and Cerqua and Letta, 2020, we use the estimators based on Eq. (2.10). These estimators capture the differences between the observed outcome, Y (binary variable accounting for the success of a Colombian exporter in 2020), and its counterfactual predictions (SUM). Figure 17 shows that the average individual treatment effect for COVID-19 is very similar when the individual treatment effect is estimated as the average of $\hat{\alpha}$ (black line, $SAM - SUM$) or as the average of $\hat{\hat{\alpha}}$ (yellow line, $Y - SUM$). As shown in Figure 17, when the interest lies in estimating the average treatment effects (by months in this case), the results based on $Y - SUM$ do not differ from those obtained by using $SUM - SAM$. We obtain similar results for the two methodologies also in terms of conditional treatment effects based on subgroups defined on firm characteristics (e.g., by industry or main export destination country).

The fact that the two estimators consistently find zero estimated effects for all 2019 exporters (and for subgroups based on the values of individual observables) during the first quarter suggests that the estimation error of both SUM and SAM, $\mathcal{E}_0$ and $\mathcal{E}_1$ respectively, goes to zero when we average the individual treatment effects across the whole distribution of 2019 exporters (or in subgroups defined by one of the possible dimension of treatment effect heterogeneity defined by observables; e.g., by industry or main export destination country).

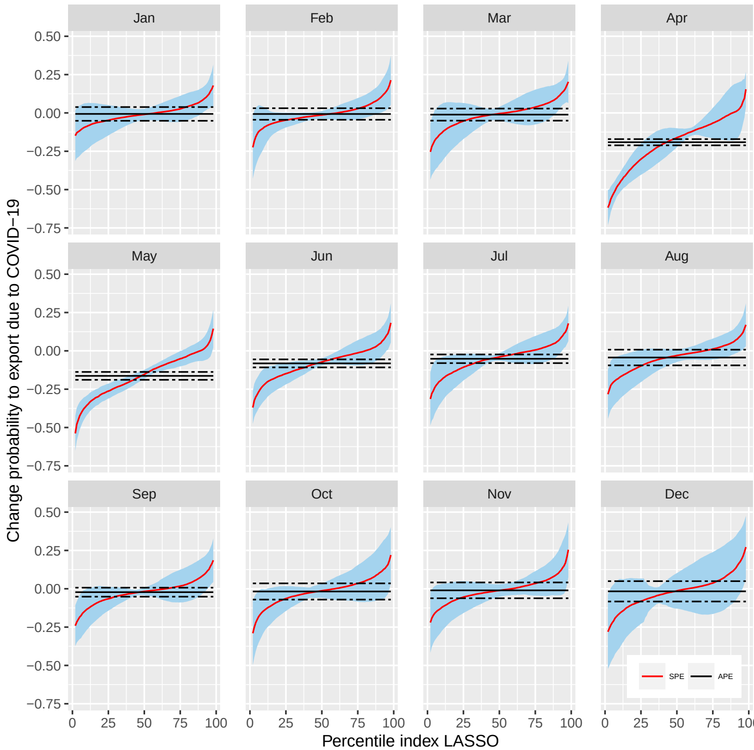
However, since our goal is to identify the main dimensions of treatment effect heterogeneity by classifying units with the highest and lowest estimated treatment effects, we need also to evaluate how well these alternative estimation strategies perform in identifying treatment effects at the extremes of the distribution of treatment effects. Figure 18 shows the average of the estimated treatment effects obtained with the two estimators for the observations whose estimated treatment effects (by using $Y - SUM$) are contained in intervals defined by two consecutive values of the estimated percentiles of $Y - SUM$.

On the one hand, the estimator based on $Y - SUM$ is also identifying significant treatment effect heterogeneity in the first quarter, suggesting that the distribution's estimation error, $\mathcal{E}_0$, is not zero on average in the tails. Moreover, the shape of the $Y - SUM$ curve is similar across quarters, suggesting that this estimation method will be prone to misclassify units when using the Sorted Effects strategy suggested above. On the other hand, in the first quarter, the shape of the $SAM - SUM$ curve is flat, showing a constant average estimated effect that is zero along the whole distribution of the $Y - SUM$ estimated effects, suggesting that by using the SAM we are able to wash out the estimation error of the SUM because $\mathcal{E}_1 = \mathcal{E}_0$.

This behavior of the estimators based on $SAM - SUM$ is consistent with the results shown in Figure 16 for the Sorted Effects analysis. Figure 19 shows that the intuition on the inadequacy of the $Y - SUM$-based estimators to identify treatment effects on the tail of the distribution is also confirmed by the Sorted Effects analysis based on this estimation strategy. When using the $Y - SUM$ individual level estimates to feed the SPE methodology, we find economically and statistically significant effects of the COVID-19 shock all along the percentile distribution in the first quarter. While it is true that, on average, $\mathcal{E}_0$ tends to be zero across all observations, these findings suggest that this is not true when we focus on specific segments of the treatment effect distribution, particularly in the tails.

Table 7 presents the classification analysis results on the sources of treatment effect heterogeneity when the $CADiff$ is estimated using the $(Y - SUM)$ approach. For all the firm characteristics we examined, we found no statistically significant difference between the most and least affected groups. This is consistent with the inability of the $Y - SUM$ approach to consistently estimate treatment effects in the tails of the α's distribution and, consequently, to identify the groups of the most affected and the least affected firms. In other words, such groups will be contaminated by the inclusion of firms wrongly classified due to the estimation error $\mathcal{E}_0$.

Figure 16: Sorted Partial Effects (SPE) and Average Partial Effects (APE) on exporter status



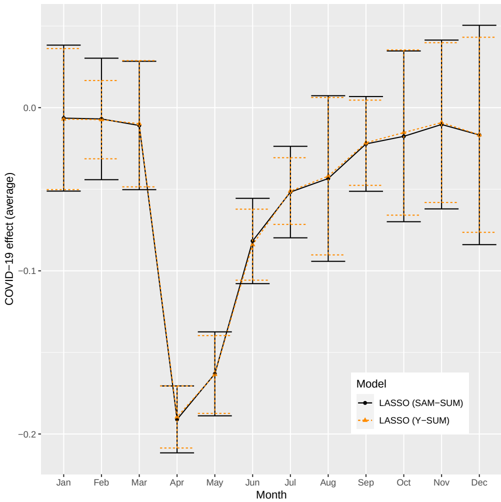
Note: Monthly Sorted Partial Effects (SPE) and Average Partial Effects (APE) of COVID-19 on Colombian firm export status. The treatment effect is calculated as a difference between SAM and SUM predictions. Standard errors were obtained with 100 bootstrap replications. Confidence intervals for a 5% significance level.

Table 6: Estimated differences in means of the estimated treatment effect and other covariates between the group of more affected and the group of less affected firms ($CADiff$) applying the classification analysis to the $SAM - SUM$ estimates

Outcome variable	(1)	(2)	(3)
TE	-0.3130^{***}	-0.3060^{***}	-0.2790^{**}
Agriculture	-0.1940		
Chemicals	-0.0057		
Manufacturing	-0.0092		
Metals	0.0134		
Special	0.0056^{***}		
Textile	0.1600^{***}		
Wood	0.0292^{***}		
Air	0.2030^{*}	0.1680^{***}	0.2040^{***}
Land	0.0340	0.0249	0.0170
Sea	-0.2360^{***}	-0.1920^{***}	-0.2200^{***}
Jan	-0.0738	-0.0766^{***}	
Feb	-0.0710	-0.0768^{***}	
Mar	-0.0751	-0.0773^{***}	
Apr	0.1860^{***}	0.1950^{***}	
May	0.1770^{***}	0.1820^{***}	
Jun	0.0754	0.0784^{***}	
Jul	0.0132	0.0159	
Aug	0.0021	0.0008	
Sep	-0.0412^{***}	-0.0406^{**}	
Oct	-0.0604^{***}	-0.0609^{**}	
Nov	-0.0723^{***}	-0.0763^{**}	
Dec	-0.0557	-0.0621^{**}	
Number of export destinations (ND)	-0.1990	-0.1640	-0.2480
Number of import origins (NO)	-1.7470	-1.9820^{***}	-2.4440^{**}
Number of exported products (NP)	0.2400	-0.2570	-0.3440
Containment Index Stringency Export	19.3600^{***}	19.5100^{***}	7.1800^{*}
Containment Index Stringency Import	19.1100^{***}	20.8000^{***}	7.2490^{***}
Value Exported (log)	-0.5110^{***}	-0.4490	-0.5700^{*}
Value Imported (log)	-1.8160^{***}	-2.2020^{***}	-2.6860^{***}
Deviation from sectoral mean		✓	✓
Deviation from monthly mean			✓

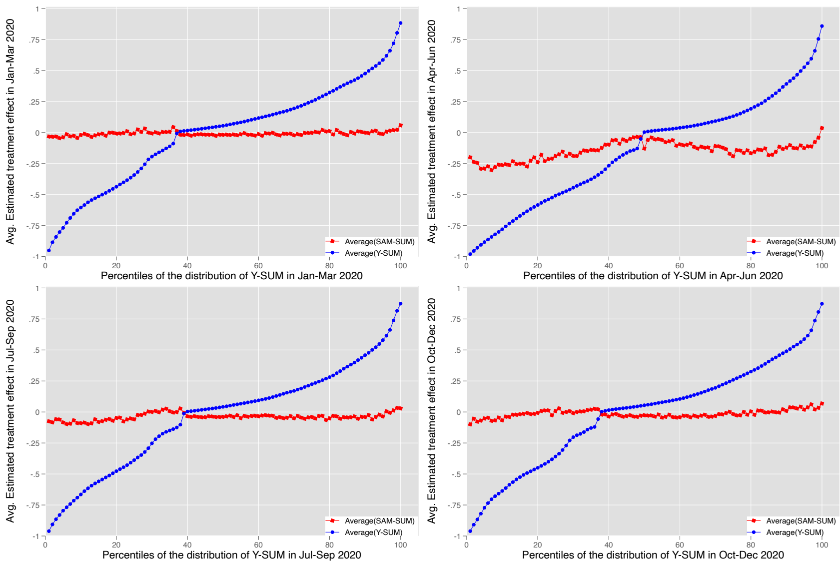
Notes: column (1) does not include sector or month variables in the regression; column (2) includes sectors in the regression; and, column (3) includes both the sector and month variables. *** means significant at 1%, ** at 5%, and * at 10%. Standard errors are obtained by bootstrapping the whole estimation process, and joint p-values are adjusted to consider the simultaneous testing of all variables. We group the 22 industries as follows. "Agriculture": Animals (01), Vegetables (02), Fats/oils (03), and Prepared Foodstuffs (04). "Chemicals": Chemical (06), and Plastics (07). "Manufacturing": Machinery (16), Vehicles (17), and Manufactured (20). "Metals": Mineral (05), Cement (13), Jewelries (14), and Metals (15). "Special": Precision Instruments (18), Arms (19), Art (21), and Special (22). "Textile": Leather (08), Textile (11), and Footwear (12). Finally, "Wood": Wood (09), and Paper (10). See Table 16 for the complete industry names.

Figure 17: ATE Comparison (SAM-SUM) vs. (Y-SUM)



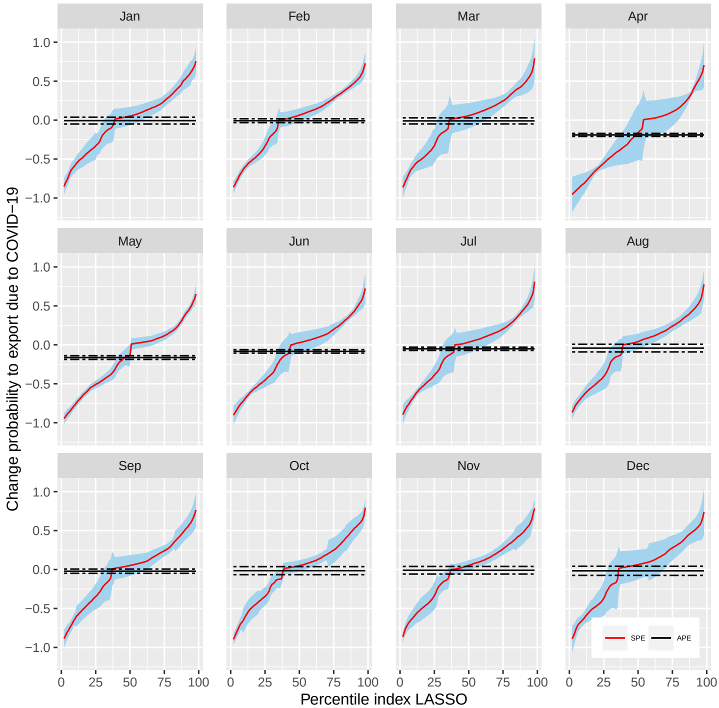
Note: Mean difference in the predicted probability of success (SAM - SUM vs. Y - SUM) by month, using Logit-LASSO predictions and (SAM vs. SUM). Standard errors were obtained with 100 bootstrap replications. Confidence intervals for a 5% significance level.

Figure 18: ATE Comparison by percentiles (SAM-SUM) vs. (Y-SUM)



Note: Estimated average treatment effects ($SAM - SUM$, red line, and $Y - SUM$, blue line) by quarter for observations contained in intervals defined by the estimated percentiles of $Y - SUM$.

Figure 19: Monthly Sorted Partial Effects (SPE) by percentiles (Y-SUM)



Note: Monthly Sorted Partial Effects (SPE) and Average Partial Effects (APE) of COVID-19 on Colombian firm export's status. TE is calculated as a difference between the observed outcome (Y) and SUM predictions. Standard errors were obtained with 100 bootstrap replications. Confidence intervals for a 5% significance level.

Table 7: Estimated differences in means of the estimated treatment effect and other covariates between the group of more affected and the group of less affected firms ($CADiff$) applying the classification analysis to the $Y-SUM$ estimates

Outcome variable	$\beta_{1,f}^{(1)}$	$\beta_{1,f}^{(2)}$	$\beta_{1,f}^{(3)}$
TE	-1.0910	-1.0930***	-1.0710
Agriculture	-0.0616		
Chemicals	-0.0192		
Manufacturing	0.0112		
Metals	0.0109		
Special	0.0059		
Textile	0.0486		
Wood	0.0041		
Air	0.0411	0.0271	0.0289
Land	0.0086	0.0062	0.0068
Sea	-0.0482	-0.0321	-0.0343
Jan	-0.0190	-0.0189	
Feb	-0.0242	-0.0237	
Mar	-0.0181	-0.0181	
Apr	0.0631	0.0630	
May	0.0620	0.0612	
Jun	0.0166	0.0167	
Jul	0.0033	0.0028	
Aug	-0.0050	-0.0053	
Sep	-0.0169	-0.0167	
Oct	-0.0216	-0.0208	
Nov	-0.0218	-0.0222	
Dec	-0.0183	-0.0181	
Number of export destinations (ND)	0.3310	0.3470	0.3306
Number of import origins (NO)	0.0350	-0.0595	-0.1077
Number of exported products (NP)	0.6050	0.4670	0.4275
Containment Index Stringency Export	-0.2280	-0.0264	0.9690
Containment Index Stringency Import	-4.2180	-4.4910	-0.0520
Value Exported (log)	-0.2700	-0.2760	-0.1800
Value Imported (log)	-0.0910	0.0296	0.0040
Deviation from sectoral mean		✓	✓
Deviation from monthly mean			✓

Notes: column 1 does not include sector or month variables in the regression; column 2 includes sector in the regression, and, column 3 includes both the sector and month variables. *** means significant at 1%, ** at 5%, * at 10%. Standard errors are obtained by bootstrapping the whole estimation process and joint p-values are adjusted to take into account the simultaneous testing of all the variables.

2.5 Concluding discussion

In this paper, we evaluate the effects of an economy-wide shock such as COVID-19 on a firm's export behaviour, identifying the most affected subpopulations and the sources of treatment effect heterogeneity in scenarios where a credible control group is unavailable, and it is difficult to define ex-ante the varying degrees of exposure to a shock for each economic agent.

From a methodological point of view, we show practitioners how to apply the generic ML tools proposed by Chernozhukov et al., 2020 to a context in which there is no control group available; we suggest how to use in-time placebo tests to check the credibility of counterfactual estimates. Furthermore, we provide evidence indicating that in a Sorted Partial Effects analysis, in which the focus lies on the tails of the distribution of the treatment effects, it is critical to correct the estimation error arising from the necessarily imperfect reconstruction of the unobservable counterfactual in order to estimate the causal effect heterogeneity. While this method is specifically designed for analyzing the heterogeneous impacts of economy-wide shocks, it is conceivable that the construction of a counterfactual for all units could extend its applicability to less extreme scenarios characterized by the presence of general equilibrium effects. In such contexts, this empirical framework facilitates the examination of potential indirect effects stemming from policies and shocks, without the prerequisite assumption of an a priori defined control group of untreated units.

Using data from the Colombian customs office, we estimate that, during 2020, on average, the COVID-19 shock decreased a firm's probability of surviving in the export market by about 15 to 20 percentage points in April and May and by approximately 5 to 8 percentage points in June and July. By analysing the estimated treatment effect distribution, we reveal that these average results hide considerable heterogeneity. For instance, in April 2020, we find that for some exporters COVID-19 decreased their survival probability by 55 percentage points. We identify the firms most and least affected by COVID-19 and we compare their characteristics by integrating the Sorted Partial Effects methodology with our causal ML approach.

Moreover, our findings underscore the role of diverse import strategies in fortifying exporters' resilience during crises like COVID-19. Integration within Global Value Chains (GVCs) —measured by both the number of countries from which exporters source and the value of imports—emerges as a pivotal role in buffering exporters against the shocks induced by unforeseen events like the global health crisis. Given these compelling findings, we advocate for a strategic policy focus, where policymakers should recognize the critical role of fostering and supporting diversified import strategies among exporters as a cornerstone of fortifying exporters' resilience. This recognition could fuel the development of policies aimed at facilitating and incentivizing broader and more diverse import networks. Measures such as trade agreement promotion, infrastructure development or regulatory reforms, have the potential to significantly contribute to the stability and adaptability of export-oriented sectors in the face of future uncertainties, thereby reinforcing economic resilience and sustainable growth.

Finally, in this paper we also demonstrate that ML methods can be applied successfully to predict firms' trade potential. We consider ML methods a promising tool to assist firms and public agencies in their export decision-making processes. The bulk of countries possesses export promotion agencies whose objective is to sustain firms' internationalization activities by lowering the costs of information acquisition (Broocks and Van Biesebroeck, 2017; Munch and Schaur, 2018).

Given that exporters' dynamics can be understood as a complex learning process dense of interdependencies (complementarity or substitutability) between products and destination markets (from the perspective of technology/knowledge, local tastes, legal requirements, and marketing and distribution costs)[18] and that ML techniques have been successfully applied to predict firm performances in such settings, we believe that an important avenue for future research is to test the effectiveness of using these techniques and firm-level data to build recommendation systems.

[18]In the core competence model by Eckel and Neary, 2010 the focus is on process innovation and productivity: each firm has a core product in which its productivity is the highest and bears adaptation costs to produce different products. In Montinari, Riccaboni, and Schiavo, 2021 (page 4), instead, product innovation and cumulative growth are central: to enlarge the portfolio of produced and exported goods, firms have to invest in R&D and " the more products a firm has, the more resources it can devote to develop additional goods." Hidalgo et al., 2007 and Jun et al., 2020 study bilateral trade at the product-country level and represent the portfolio of goods shipped by countries as a network, the so-called product space. They define a measure of relatedness between products based on co-exporting patterns to capture common capabilities and knowledge flow between products. Finally, Morales, Sheu, and Zahler (2019) find that the entry costs a firm has to bear to start exporting to a new market depend on the similarity of the new destination with respect to those of the portfolio of markets already reached by the firm (in terms of geographic location, language, and income per capita) and profit shocks in a market affects firms' exports in other markets.

These systems could help firms identify their latent comparative advantages and provide export diversification and differentiation recommendations.

Chapter 3

Exports' Survival in New Markets: A firm-level export recommendation model.

3.1 Introduction

Export promotion activities (EPA) are essential in supporting firms' efforts to enter and succeed in international markets. These initiatives can take various forms, ranging from financial support, such as subsidies and credit lines, to information dissemination. Information-sharing efforts often include organizing trade fairs, facilitating connections with importers, and providing insights on foreign demand and product standards to reduce trade frictions. The primary aim of these policies is to lower the fixed costs associated with acquiring market information, thus enabling domestic firms to explore new markets or products more effectively.

This paper introduces a novel approach within the realm of EPA by proposing a firm-product market recommendation tool that leverages past export market entries in Colombia, modeled using Random Forest

[0]This chapter is based on my solo-authored work, 'Exports' Survival in New Markets: A firm-level export recommendation model'

techniques. Unlike traditional EPA efforts, which broadly disseminate market information, this tool offers tailored recommendations by analyzing the historical entry and survival experiences of exporters in specific product-markets. This approach diverges from conventional EPA methods, which often overlook the importance of survival outcomes in market entry strategies, as highlighted by Esteve-Perez, Requena-Silvente, and Pallardo-Lopez, 2013 (page 22), who noted that "export promotion policies prompting entry into export markets without taking into account survival may be flawed." It addresses how firms can benefit from these data-driven insights when selecting new export destinations, ultimately enhancing their understanding of demand and supply dynamics. The broader implications for export product growth and economic development in Colombia are also considered.

The development of this recommendation tool is particularly relevant given the persistent challenges faced by exporters, especially in developing countries. Alvarez, Faruq, and López, 2013 indicate that firms often under-invest in entering new markets due to their inability to fully internalize the benefits of exporting, such as positive spillover effects. This investment below the socially optimal levels highlights the need for targeted government intervention to facilitate market access. Moreover, exporters frequently encounter difficulties in accessing comprehensive market data (Volpe Martincus and Carballo, 2010), complicating decision-making processes when venturing into new markets. These challenges are exacerbated in developing nations, where international market structures may be underdeveloped, and export promotion budgets are often limited (Olarreaga, Sperlich, and Trachsel, 2020). For instance, in Colombia, large and volatile exit patterns among exporters have been observed, particularly during global trade shocks (Dueñas et al., 2021).

Information uncertainty plays a crucial role in the premature termination of many export ventures, as highlighted by Besedeš and Prusa, 2011. This uncertainty underscores the importance of effective information dissemination and tailored support. While existing literature has focused on country-level export recommendations (e.g., Hidalgo et al., 2007; Che, 2020; Decreux and Spies, 2016), there has been a notable gap in the development of firm-level market recommendation tools.

This paper makes significant contributions to two key areas of the literature. First, it introduces the first firm-product-market level recommendation system designed specifically for exporters, representing a novel approach in the literature. Second, it provides empirical evidence on the positive effects of such tailored EPA on export outcomes. By simulating what would have happened if Colombian exporters had followed the model's recommendations for new market entries in 2018, the paper finds that these exporters would have experienced a 5 percentage point higher survival rate compared to those choosing different markets. Furthermore, products exported to destinations where at least one exporter coincidentally aligned with the recommended strategy exhibited significantly larger growth compared to those where no exporters followed any similar recommendation.

These findings align with the broader literature on the positive impacts of EPA on entering new markets (Volpe Martincus and Carballo, 2008; Volpe Martincus and Carballo, 2010; Broocks and Van Biesebroeck, 2017), increasing GDP per capita (Olarreaga, Sperlich, and Trachsel, 2020), and boosting total exports (Munch and Schaur, 2018). Additionally, this paper presents preliminary evidence that aligns with prior studies, such as Békés and Muraközy, 2012 and Esteve-Pérez, 2021, regarding the prevalence of exporters participating in temporary trade spells. The findings suggest that exporters incur temporary trade as a way of experimentation and to resolve knowledge uncertainty or incomplete information. This perspective indicates that exporters may exhibit myopic decision-making processes due to inadequate market information when making strategic choices about new market entries, potentially leading to sub-optimal market entry decisions.

Overall, this paper provides Colombian exporters with a strategic tool to improve survival probabilities and accelerate product export growth, which could have significant implications for Colombia's economic development and international trade competitiveness.

The paper is organized as follows: Section 3.2 presents the Colombian firm-level data, the variables used in the analysis, and the export survival dynamics in Colombia. Section 3.3 outlines the empirical strategy, Section 3.4 reports the main estimation results, and Section 3.5 summarizes the findings and discusses the implications and limitations of the analysis.

3.2 Data & Export Survival in Colombia

3.2.1 Data

To create the market-recommendation model and assess its impact on export survival and product growth I will use Colombian export transactions data reported by the Colombian Customs Office (DIAN) for the years 2006-2020. This dataset tracks each export performed from Colombia to the rest of the world. It includes variables such as the date, the exporter identifier (NIT), a 6-digit Harmonized System code (HS)[1] that describes the product, the Colombian department from which the product originates, means of transportation for the export, country of destination (market), free on board value (FOB), as well as the insurance and the freight cost in US dollars (CIF).

I remove all transactions related to re-exports of products elaborated in other countries and all product-market combinations lower than $\$1000$ at firm-level in a given year, following the approach in Fugazza and Molina, 2016. As a result, I end up with $883,568$ firm-product-market spell reports in the period 2006-2020 ($45,543$ different exporters).

In line with Alvarez, Faruq, and López, 2013, I extract various features from the original data set, which I categorize into firm characteristics (internal information) and market characteristics (external information). The internal features include the number of products (NP), the number of export destinations (ND), and the number of varieties (NV) a firm serves in year t. Moreover, the existing literature has firmly established the presence of interconnected (interdependent) markets, as demonstrated and modeled by Chaney, 2014, Morales, Sheu, and Zahler, 2019, Alfaro-Urena et al., 2023, and Albornoz et al., 2023. Therefore, I incorporate a set of variables that capture the weighted accumulated experience by destination. In particular, I measure previous experience as the firm-weigh of previous exported value by destination against the total previous exported value across all destinations. This inclusion enables the model to account for the impact of previous experiences and cross-country complementarities in firm exports. Exporting to certain countries positively affects the probabilities of firms engaging in exporting to additional countries as found in Alfaro-Urena et al., 2023. This approach

[1] The HS is updated every five years affecting some codes in this analysis that suffered changes in the years 2007, 2012, and 2017. To keep the products affected in the analysis I concord the product codes affected by updates, following the methodology proposed by Van Beveren, Bernard, and Vandenbussche, 2012.

allows to not only capture whether an exporter had experience in certain destinations, but also the relative importance of those destinations in their overall export portfolio. The inclusion of these variables holds significant importance, given the robust empirical evidence demonstrating that prior export experience plays a pivotal role increasing both the probability of engaging in export activities (Alvarez, Faruq, and López, 2013) and the likelihood of export survival (Albornoz, Fanelli, and Hallak, 2016, Esteve-Pérez, 2021). These findings highlight the essential nature of considering previous export experience as a key determinant in understanding and predicting export dynamics and export survival. Furthermore, to account for other learning and maturation patterns, I design additional firm-level variables. Specifically, I consider whether it is the first time the firm is exporting a particular product as a dummy variable to signal any product entry. I also create a variable accounting for the, firm-level, previous value of exports for that product and for the number of years that a firm was exporting a given product until period t-1, as well as the number of years exporting in the past in the same product but other markets, and a different product in different markets.

In relation to market information, I incorporate various variables to capture the size of each market and product. Firstly, I include the number of firms selling a particular product, in a given market, as well as the number of firms selling the same product in a specific product-market combination. Additionally, I include the lagged versions of these variables to account for any time-dependent dynamics. Moreover, I consider the number of products available in a specific market to proxy its size, reflecting the range of product offerings within that market. Similarly, I include the number of different destinations associated with each product to provide insights into the product's reach and diversification across markets. Lastly, to assess the significance of a firm's product within a market, I incorporate a variable that measures the share of the firm's product-market spell over the total product sales in this market. This variable explains the relative prominence and competitiveness of the firm's product within the overall market. To account for any persistence in market share dynamics, I also include the lagged version of the product share variable.

To create the survival outcome variable it is important to note that exporters can experience "export amnesia" if they remain inactive for a period of time, during which they may be unaware of demand shocks in the market. The evidence provided by Roberts and Tybout, 1997 suggests that information depreciates rapidly for Colombian exporters who reenter the market after a 2-year period of inactivity. Additionally, Besedeš and Prusa, 2011 find that new export relationships in developing countries rarely persist beyond two years. These empirical findings support the adoption of a conservative definition of survival by employing a two year timeframe for the outcome variable. Furthermore, to address the potential influence of export inactivity on the exporter's knowledge and response to market conditions, a dummy variable is included to capture whether the exporter has been active in the last two consecutive years, regardless of the specific product or market.[2]

3.2.2 Export Survival

In order to establish a clear definition of export survival, it is important to define the concept of export incursion. Building upon the framework of Albornoz, Fanelli, and Hallak, 2016, I define market incursion as the initial entry of an exporter into a specific market (M). Subsequent re-entries into the same market are not considered incursions. Following these lines, export survival within a specific product-market (PM) occurs when the firm continues to export in that PM two years after the initial incursion. The model in Section 3.3 aims to estimate the probability of export survival two years after the market incursion.

Figure 20 presents empirical findings regarding the duration and number of attempts necessary for an exporter to achieve initial survival in a market (left panel), in a product (center panel), and in a product-market combination (right panel). Notably, the observed pattern exhibits some similarities across all three types of survival.[3] It is observed that firms that successfully survive typically undertake between 3 and 7 market incursions and experiment with 4 to 9 different product varieties. However, a significant proportion of firms (represented by the light blue color) do not survive during the analyzed time period, and the median exporter not surviving ventures into only one market, and tries only a single product. It is important to note that the empirical visualization only includes new exporters from 2008 with an annual FOB value in the variety exceeding \$1,000.[4] Therefore, the findings underscore the significant time

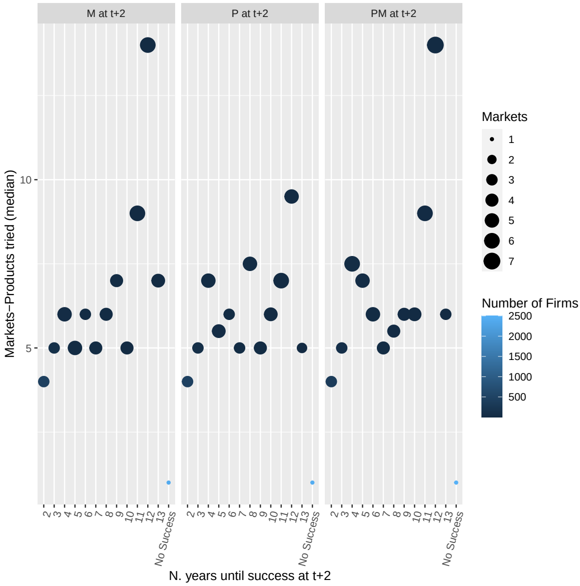
[2]For a summary of all features see Table 17 in Appendix C.1.

[3]Survival at the PM level is of particular interest as the initial incursion occurs at the market level, while the exporter employs one or more products to enter a new market. Thus, the focus lies on the duration of the variety spell survival.

[4]Fugazza and Molina, 2016 examined the determinants of export survival using a similar selection criterion.

and experimentation required for export survival, involving years of trial and the exploration of new products and markets.

Figure 20: N. of incursions at PM, M, and years for export survival



Note: N. of incursions at PM, M, and years needed to survive exporting either at (M), (P), or (PM) in $t+2$, in Colombia (2008-2019) - Only new exporters in 2008, and more than $1,000 FOB value (as in Fugazza and Molina, 2016).

Table 8 presents data on the annual number of market incursions that occurred in Colombia between 2008 and 2019.[5] In 2019, there were approximately 2,500 fewer incursions and around 900 fewer exporters involved in these incursions compared to 2008. The market-survival rate of market incursions exhibits variation over the years, ranging from 18% in 2008 to a peak of 24% achieved by firms entering new markets in 2010. On average, the product-market survival rate of market incursions stands at 20%. Notably, the product-market survival rate of market incursions surpasses the product-market survival rate of product incur-

[5]To prevent biased results stemming from observations considered as market incursions due simply to the initial data years of 2006 and 2007, the analysis excludes these years, although exporter behavior during these years is also observed.

sions (as indicated in the last column of Table 8) for each year, with an average survival rate of 16%.

Table 8: Market incursions, value, and survival rates

Year	Export Value (millions US$)	# Firms	Export Value Mean (in thousands US$)	# Incursions at M	# Firms (incursions at M)	Value per incursion at M (mean in thousands US$)	Survival Rate of M at PM (t+2)	Survival Rate of P at PM (t+2)
2008	24353	9813	395	12300	5760	158	0.18	0.12
2009	22767	9549	396	10614	5299	112	0.20	0.15
2010	28715	8444	529	9760	4689	572	0.24	0.20
2011	37434	8579	655	9590	4730	386	0.20	0.17
2012	37978	8768	651	9571	4765	302	0.22	0.17
2013	36481	8980	633	9468	4791	219	0.19	0.16
2014	33794	9005	585	9374	4788	202	0.18	0.15
2015	25747	9223	442	10709	5100	167	0.22	0.17
2016	23477	9326	404	9762	5031	117	0.19	0.15
2017	28656	9459	482	9869	5023	182	0.19	0.16
2018	30507	9688	505	10091	5142	138	0.19	0.15
2019	28246	9574	458	9764	4826	133	0.21	0.17

Note: Colombia Exports, 2008-2019. Only yearly exports of more than $1,000 FOB value. Note: M, P, and PM stand for Market, Product, and Product-Market respectively.

Table 9 presents empirical findings regarding the duration of export spells for exporters that entered new markets in 2008, aggregating a total of 12,300 market incursions as demonstrated in Table 8. Interestingly, more than 66% of new export ventures to different destinations conclude within the first year of entry. This rate is slightly higher than the corresponding figure for Spanish exporters, as reported in Esteve-Pérez, 2021, where 60% of new product-country export ventures terminated during their inaugural year. Besedeš and Prusa, 2011 also finds evidence that export survival for developing economies is shorter than that for developed economies. This disparity could possibly be influenced by Colombia's comparatively lower level of development in international trade. For instance, Colombian exports accounted for approximately 0.2% of global trade in 2021, whereas Spanish exports constituted around 1.7% of global trade in the same year (World Bank, 2021).[6] Similarly, Colombia's trade-to-GDP ratio stood at 48% in 2022, whereas Spain's ratio was significantly higher at 81% (World Bank, 2022).[7] These indicators highlight the disparities in trade development between Colombia and Spain, potentially influencing the duration of export ventures in their respective markets.

Despite the fact that the majority of market incursions in Colombia terminate within the first year, they exhibit notably higher survival rates than product entries, as shown in the last two columns of Table 8. This

[6] The global trade data is collected from the WITS dataset from the World Bank.

[7] Trade to GDP indicators are collected from the World Bank dataset.

Table 9: Number of incursions, at Market, by duration

Duration	1	2	3	4	5	6	7	8	9	10	11	12+
# Incursions (M)	8184	1902	663	394	251	203	143	90	78	57	43	292

Note: Number of incursions, at Market, by duration. Sample from Colombian export spells that start exporting in a new market in 2008. Only yearly exports of more than $1,000 FOB value.

finding highlights the potential for enhancing export survival through a market recommendation system, given that firms entering new markets tend to have better chances of survival than those introducing new products. This suggests that exporters face fewer barriers to sustaining market entries compared to launching new products, underscoring the value of focusing on market-level export promotion efforts.

The emphasis of this paper, therefore, is on export promotion at the market level rather than the product level. Recommending new markets, rather than new products, not only aligns with the empirical evidence on survival rates but also has a practical rationale. The dataset used in this analysis classifies products at the 6-digit HS code level, which can aggregate a wide range of heterogeneous items. For many firms, especially those in developing economies, it may be more feasible to explore new destinations than to switch to an entirely new product line without prior expertise or experience.

This focus on market entry aligns with broader concerns in the literature. Studies by Besedeš and Prusa, 2006a, Besedeš and Prusa, 2011, Hess and Persson, n.d., and Nitsch, 2009 highlight that trade relationships at the country level are often short-lived. Similarly, Békés and Muraközy, 2012, and Esteve-Perez, Requena-Silvente, and Pallardo-Lopez, 2013 find that this short-term nature extends to trade relationships at the firm level. The frequent interruption of these relationships has been shown to negatively impact economic growth, as demonstrated by Fugazza and Molina, 2016. By addressing these challenges through targeted market recommendations, export promotion strategies can help mitigate the risks associated with premature market exits and support more sustainable growth trajectories.

3.3 Empirical Strategy

In this section I detail the benchmark model[8] that will estimate the exporter status after trying a new market that will form the basis to later construct the recommendation issuer model. Following Alvarez, Faruq, and López, 2013 I define:

$$Y_{ipmt+2} = \alpha Z'_{ipmt} + \beta W'_{ipmt} + \delta_t + \delta_p + \delta_m + \varepsilon_{ipmt} \tag{3.1}$$

where,
Y takes value 1 if firm i survived exporting a given product p in market m in year $t+2$ after a market entry in m at period t,[9] 0 otherwise.
Z' contains the variables that are internal to the firm $\rightarrow$ firm-specific.
W' contains product and market information variables $\rightarrow$ firm-external.
δ'_t, δ_p, δ_m : time, product, and market FE.

In order to utilize this model for prediction purposes, I adopt the standard procedure commonly employed in the field of Machine Learning (ML). This involves conducting out-of-sample predictions and validating the model's performance. Considering the temporal nature of the data, I employ a training and testing approach. The model is fitted using historical observations (training phase) and subsequently used to make predictions on future observations (testing phase). It is important to note that a restriction is imposed, whereby only products included in the training set are allowed in the testing set. This restriction ensures that the model does not make predictions for entirely new products, as the focus lies in predicting survival in new markets. As an additional assumption in the model, I treat each observation as if it were generated by a distinct exporter, even though this may not necessarily reflect the reality. The rationale behind this assumption is that the model calculates a survival probability by leveraging the accumulated knowledge from comparable market trials in the past, regardless of the exporter's identity. Furthermore, the model accounts for the impossibility of a firm attempting entry into the same market more than once.

To mitigate the absence of firm fixed effects in the model, it is possible to introduce control variables based on firm productivity. Unfortunately,

[8]The model only includes exporters trying a new market with at least $1000 FOB value in the variety tried this year.

[9]The outcome variable will be also evaluated as the survival at product level, at the market level, and at exporter level respectively for some exercises.

the available dataset does not provide any measure of productivity, capital intensity, or firm size. However, following Alvarez, Faruq, and López, 2013, I use the lagged value of exports as a proxy for productivity.[10] This approach is motivated by the observation that larger firms tend to exhibit higher levels of productivity, and firm size is often associated with the costs of export entry. By incorporating the lagged export value as a proxy variable, I aim to capture the potential effects of productivity and firm size on the model's outcomes.[11]

3.3.1 Model Selection

To give an accurate market recommendation the model has to perform good at predicting correctly the exporter survival cases, called "sensitivity" indicator in the literature, in a new market, rather than being able to correctly predict the cases when an exporter does not survive in a new market, called "specificity" indicator in the literature. However, to avoid that the model only predicts the survival category (1's) one should also look at the "accuracy" measure which gives an intuition on how well the model does in predicting the two type of outcomes. Therefore, in this context, it is recommended to focus on accuracy and sensitivity as measures of goodness of fit (GoF).

In the present study, I employ a Logit model as the natural benchmark within the specified context. However, given the intricacies of this highly-dimensional dataset and the complex, non-linear relationships inherent in the features influencing the variable of interest, I seek to enrich the analysis by comparing the out-of-sample prediction performance of the Logit model with that of a standard Random Forest, as introduced by Breiman, 2001. To provide a clearer perspective, the main distinction between these models using the notation employed by James et al., 2013, is that in the linear regression framework, the functional form is expressed as follows:

$$f(X) = \beta_0 + \sum_{j=1}^{p} X_j \beta_j \tag{3.2}$$

[10]Firm-specific variables such as productivity, as discussed in Roberts and Tybout, 1997 and Bernard and Jensen, 2004, are absent in the dataset utilized for this study. Hence, in line with Alvarez, Faruq, and López, 2013, productivity is approximated using the lagged value of exports.

[11]See Appendix C.1 for a deeper explanation regarding the variables introduced in the model.

, where j represents a feature that predicts the export survival probability. Conversely, the Random Forest (RF) model, and regression tree methods in general, embrace a different functional form:

$$f(X) = \sum_{m=1}^{M} c_m \cdot 1_{X \in R_m} \tag{3.3}$$

, where $R_1, ..., R_M$ represents the partition of the feature space. The RF model aggregates the predictions from multiple random decision trees, each delineating distinct regions in the feature space, thereby capturing non-linearities and interactions that may elude a conventional linear framework.

Existing evidence indicates that machine learning (ML) models, such as Random Forest (RF), offer superior out-of-sample prediction performance compared to traditional methods in international trade analysis, as demonstrated by Dueñas et al., 2021. In contrast to the Logit model, Random Forest exhibits a more opaque nature, often referred to as a "black box." Notably, it does not provide estimated signs for individual independent variables included in the model. However, Random Forest offers valuable insights in the form of variable importance rankings, which indicate the predictive power of each variable included in the model. This ranking serves as a measure of variable importance.

Unlike conventional hazard models used for analyzing duration and survival in trade, such as those by (Esteve-Perez, Requena-Silvente, and Pallardo-Lopez, 2013, Besedeš and Prusa, 2006a, Besedeš and Prusa, 2006b, Fugazza and Molina, 2016), Random Forest can better accommodate non-linearities and complex interactions among variables. While traditional hazard models often rely on linear assumptions, limiting their ability to capture intricate patterns, the Random Forest approach provides a more flexible framework for modeling complex relationships and enhancing predictive accuracy.

Table 10 presents the measures of goodness of fit (GoF) for various prediction types in the context of export survival. The first column represents the GoF values for predicting the exporter status following the exploration of a new market. The second column displays the GoF measures for predicting the firm-product status after venturing into a new market. The third column showcases the GoF values for predicting the firm-market status subsequent to entering a new market. Lastly, the fourth column shows the GoF measures for predicting the firm-product-market status following the firm's expansion into a new market. Considering the objective of achieving a model with superior accuracy and sensitivity measures, the most suitable choice would be Random Forest (RF) as the preferred model given that High-Dimensional-FE Logit always presents lower accuracy and sensitivity than RF.

Table 10: Variable Importance - Export Survival after a market entry

Continue exporting, after market entry, in period t+2 with the same...				
Var. Importance	**id**	**Product**	**Market**	**Product-Market**
1	NM	NM	NV	Value in PM (t)
2	NV	Acc. Value in P (t-1)	% Firm's PM portfolio (t)	NM
3	% Firm's PM portfolio (t)	N. years in SP-OM (t-1)	NP	NV
4	N. years in OP-OM (t-1)	N. years in P (t-1)	NM	% Firm's PM portfolio (t)
5	N. years in SP-OM (t-1)	FP	N. firms in P (t-1)	N. firms in P (t-1)
6	NP	NV	N. firms in P (t)	N. markets in P (t)
7	Small firm (Q1)	% Firm's PM portfolio (t)	N. products in M (t)	N. firms in PM (t)
8	Acc. Value in P (t-1)	Large firm (Q4)	N. firms in M (t)	Market Share in PM (t)
9	N. years in P (t-1)	N. firms in PM (t-1)	Value in PM (t)	N. firms in PM (t-1)
10	Experienced exporter	Value in PM (t)	N. firms in M (t-1)	N. markets in P (t)
Random Forest - RF				
Accuracy	78%	79%	70%	74%
Sensitivity (1's)	79%	82%	74%	78%
Specificity (0's)	78%	76%	61%	58%
Logit High-Dimensional-FE				
Accuracy	75%	77%	64%	64%
Sensitivity (1's)	75%	77%	62%	63%
Specificity (0's)	76%	77%	68%	71%

Note: Variable Importance rank and - exporter status (at firm, firm-product, firm-market, or firm-product-market level respectively) in period t+2 after trying a new destination markets. Prediction (2017), train (2010-2016). Variables highlighted in blue color contain market-related information while variables in black contain firm-related information.

Interestingly, Table 10 also showcases the variable importance ranking obtained from predicting each dimension of export survival. An intriguing observation is that variables related to market information (highlighted in blue) – which could be argued as external to the exporter – demonstrate greater relevance in predicting firm-market survival and firm-product-market survival, compared to predicting firm survival or firm-product survival, where variables containing firm-specific information take more relevance.[12]

[12] The GoF measures and the amount of market-related variables among the Top-10 most relevant factors are robust to the iterative removal of 10% variables multiple times as Table 18 in Appendix C.3 shows.

3.3.2 Recommendation

Table 10 presents empirical evidence highlighting the significance of employing Random Forest (RF) as a reliable method for accurately estimating the probabilities of survival in product-destination spells following market incursions. As previously discussed, the model provides precise survival probabilities for the specific product-market where the firm initiated the market incursion. However, it is unable to estimate the probabilities of survival in other potential new destinations. To develop a recommendation system based on this model, a minor modification is necessary. This modification involves restricting the model's training phase to focus on only one of the top 20 destinations. This restriction ensures an adequate quantity of export trials before providing recommendations for specific markets. Subsequently, a separate model is fitted for each of these m destinations (where m= $1, \ldots, 20$). Following the model fit for a particular market (m), predictions are made in the testing phase, one year after, for exporters that have never targeted destination m. The rationale behind this modification is to achieve an exceptionally accurate model fit that can be extrapolated to similar firms that have not yet explored this specific market. This approach enables the estimation of accurate probabilities of survival in different new markets, equivalent to estimating a counterfactual scenario.

3.4 Results

Using the most accurate model presented in Section 3.3.2 I build the market recommendation tool, that provides the estimated product-market probability of survival in each destination, out of the top 20 main destinations, that is new to the exporter. When the probability of survival associated to product, in a new market, exceeds 50%, the destination is classified as recommended, and it is attached to the previous combination of exporter product-market pair. Consequently, a firm's product-market spell can be associated with none, one, or multiple destinations recommended by the model.

The recommendation system generates predictions based on data available up to 2017, enabling an analysis of whether any exporter decision in 2018 coincidentally aligned with the recommendations made in 2017, despite not having access to these suggestions. This allows me to observe how firms performed when entering new markets in 2018, and whether their choices unintentionally matched the model's recommendations. Table 11 shows that exporters who aligned with the model's recommended market in 2018 exhibited a 5 percentage point higher probability of survival compared to those who exported the same product in 2017 but chose a different market in 2018. This difference becomes even more pronounced, rising to 15 percentage points, when compared to exporters who attempted any new market entry in 2018, not aligned with the recommendation. Although exporters were unaware of the model's specific recommendations for 2017, some exporters may have followed similar strategies due to their own market analysis or unrelated factors, making this analysis a valuable back-testing exercise to assess the model's validity. Additionally, not every export spell is linked to the recommendation. Some spells did not receive any recommendation if their predicted probability of survival was consistently below 50% in of the 20 potential destination markets. These firms, with lower chances of success in new markets, might have ventured into other destinations without the benefit of a guided strategy.

The data clearly shows that exporters who entered markets aligned with the recommendations achieved significantly higher survival rates than those who ventured into new markets without aligning to the model's guidance. This suggests that adherence to the recommended strategy could have substantial benefits in improving firms' long-term export success, particularly when firms lack complete market information or make sub-optimal decisions when selecting new market entry points.

These findings align with prior studies (e.g., Békés and Muraközy, 2012; Esteve-Pérez, 2021)), which suggest that exporters engage in temporary trade as a way to experiment and resolve knowledge uncertainty or incomplete information. This paper further indicates that such uncertainty may lead to sub-optimal market-entry decisions.

Table 11: Export Survival after adhering to a recommendation

Prob. Survival for each type of export in 2018.	
1) $E(Survival_{18-20} = 1 \mid FM_{18} = 1)$	19 %
2) $E(Survival_{18-20} = 1 \mid ExpProd_{17} = 1, FM_{18} = 1)$	29 %
3) $E(Survival_{18-20} = 1 \mid ExpProd_{17} = 1, FM_{18} = 1, Recommendation_{18} = 1)$	34 %

Note: Probabilities of survival at product-market (PM) in 2018 under three different conditions: (1) conditioning on trying a new market (FM), (2) conditioning on exporting P in 2017 and trying a new market in 2018, and (3) conditioning on exporting P in 2017 and adopting any market recommendation in 2018.

Having established that exporters who coincidentally aligned to the recommendation achieved higher survival rates compared to those who did not, it is crucial to consider the potential presence of compensation effects among Colombian exporters, which may not necessarily yield overall benefits for Colombia. For instance, in the context of Colombia's significant flower exports, exporters targeting markets where other exporters have already established a presence can potentially increase their market share, but this may come at the expense of reducing the market share of incumbent exporters. Such a situation can result in a zero-sum or even a negative-sum game. However, Figure 21 provides empirical evidence indicating that product-market groups where at least one firm unintentionally matched the model's recommendations tend to exhibit higher average growth rates in their respective destinations compared to products-markets where no exporters matched the recommendation,[13] indicating that following the model recommendation to target new markets can be a successful form to increase survival rates, while expanding product growth, for Colombia.

Equation 3.4 empirically tests the findings depicted in Figure 21 by conducting a regression analysis:

$$\text{Growth}_{p,d}^{18/17} = \alpha_0 + \alpha_1 \cdot \text{I}(\text{AlignRecomm}_{p,d}^{18}) + \varepsilon_{p,m} \tag{3.4}$$

, where the dependent variable is the product export growth at aggregated level from year 2017 to 2018, and the variable of interest is the *AlignRecomm* dummy variable. *AlignRecomm* takes value 1 for destinations where at least one exporter coincidentally matched the market recommendation (i.e., exporters who were already exporting the same product in 2017 but made an incursion in their associated recommended market, in 2018).

Table 12 corroborates the findings from Figure 21, demonstrating that products exported to destinations where at least one exporter aligned with the market recommendation show, on average, a growth rate nearly 30 percentage points higher than the same products in destinations where no exporter matched any recommendation. This estimation only includes products that appear in both types of destinations: those where the recommendation was matched by at least one exporter and those where no firm aligned to the recommendation during their market entry.

[13] The definition for product-destination bundles where no firm followed the recommendation includes products that have not received any recommendation from the model.

Figure 21: Product export growth from 2017 to 2018 after coincidentally matching a market recommendation in 2018.

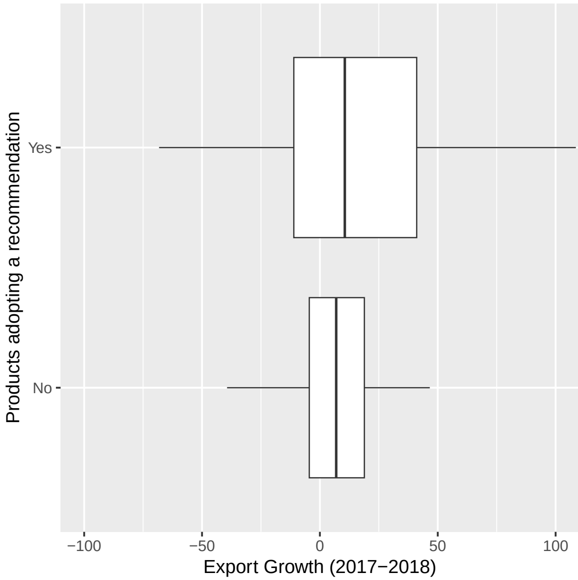


Table 12: Impact of market recommendation on product growth - OLS

	Product Export Growth (2018 vs. 2017)
Constant	8.98
	(9.78)
Dest. Aligned Recomm. (any)	29.57*
	(13.84)
Observations	176
R^2	0.026
Adjusted R^2	0.019
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*	

Note: This regression estimates the impact of coincidentally aligning with a market recommendation, in 2018, on product growth for Colombian exporters.

A total of 88 distinct products at the 6-digit HS code level meet these criteria. The sample for the regression analysis thus consists of these 88 products, each divided into two groups. The first group includes all destinations where at least one exporter matched the recommendation (*AlignRecomm*=1), while the second group encompasses destinations where no exporters aligned with the recommendation (*AlignRecomm*=0). This separation allows for a clear comparison of growth outcomes between destinations where the recommendation was matched and those where it was not, providing robust evidence of the significant impact that aligning with market recommendations can have on export performance.

The estimated effect of following the market recommendation may represent a lower bound, given the limited number of product-market spells within each of the 88 (*AlignRecomm*=1) groups. However, this result should be interpreted as a strong correlation rather than a definitive causal relationship. This limitation arises from the fact that the regression analysis does not account for some relevant factors explaining product growth, such as the productivity, or employment of firms operating within a specific product-destination group, which are not provided in the available dataset. It is worth noting that to check if this correlation could be potentially be turned into a causal link, one can further explore techniques such as a Differences in Differences Event-Study via two-way fixed effect Ordinary Least Squares (TWFE OLS) following Miller, 2023:[14]

$$\text{Growth}_{p,t} = \sum_{k=2014}^{2019} \beta_k \times \text{AlignRecomm}_{p,k-2018} + \phi_p + \gamma_t + \epsilon_{p,t} \quad (3.5)$$

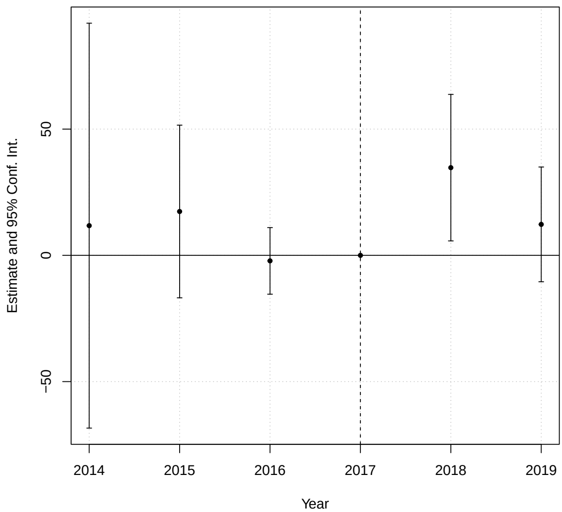
, where I interact the product-market where at least one firm adopted a new market coinciding with the recommendation issued for firm-products in 2018 ($AlignRecomm$), with a year dummy variable β_k. $Growth_{pt}$ represents the year-to-year (t) export product (p) growth from 2014 to 2019, using 2017 as the base year as it coincides with the year before issuing the market prediction by the algorithm. This equation also controls for product and time fixed effects. The coefficients after the event has occurred (β_k for $k \geqslant 2018$) indicate the product growth effects of adopting a new market that coincides with the recommendation issued in 2018.

The results coming from the event-study graph depicted in Figure 22 show that start exporting to the recommended destinations before the recommendation was done is not statistically significant when explaining product growth in the years before the recommendation is issued. This result reinforces the parallel trend assumption as well as not having anticipation effects, meaning that products where firms aligned with a

[14]The Event Study specification will reduce the number of observations as it requires to compare the same product under treatment and no treatment scenario during more years.

new recommended market in 2018 did not change their growth dynamic in anticipation to this change. However, in the year following the recommendation (2018), a compelling case emerges for a potential causal link, as evidenced by a substantial short-run increase in product growth by nearly 34 percentage points. This growth can be attributed to a strategic shift in export destination taken from non-myopic exporters, that is aligned with the export recommendations generated by the algorithm.

Figure 22: Impact of market recommendation on product growth - Event Study



Note: This regression estimates the impact of following market recommendations, in 2018, on product growth for Colombian exporters. The regression is equivalent to an Event Study from 2014-2019.

Alternatively, one can also follow Rosenbaum and Rubin, 1983 using a Propensity Score Matching (PSM) model as a tool to address the fundamental problem, not observing at the same time the treatment and control outcomes, in causal analyses. The results coming from the Event Study, in this case, could be interpreted as completely equivalent to the ones obtained following a Propensity Score Matching strategy because

PSM conducts a product-to-product (exact matching) comparison between the treated, product-market ($p - i$) groups at coinciding with recommended markets, and control units, product-market ($p - j$, where $j \neq i$) groups where there are no exporters trying new markets coinciding with the recommendation. Table 13 suggests the same estimated result between PSM and the Event Study in Figure 22. PSM proves to be more suitable than the OLS regression analysis presented in Table 12, primarily due to the inherent strength of PSM in rendering comparisons between units that are comparable. PSM achieves this by exact matching units at the product level, using the Harmonized System's six-digit classification (6-digit HS). This method juxtaposes the growth of identical product in destinations where no firm matched any recommendation (control) with those destinations (of the same product) where at least one exporter matched some form of destination recommendation (treated). Table 13 presents the results from PSM and suggests that aligning with the recommendations is associated with an estimated increase of almost 34 percentage points in product growth.[15]

Therefore, the benefits derived from matching the market recommendation can boost export survival rates as well as product growth for Colombia, which is in line with Besedeš and Prusa, 2011 that show that developing countries would experience higher export growth if they were able to improve their performance with respect to survival rates.

Notably, the 81 products outlined in Table 13 comprise 5.3% of the total number of products exported from Colombia to new markets in 2018. However, these products hold significant importance for Colombia, as they contribute to 80% of the total export value of products exploring new markets in the same year.[16] It is important to note that this study is constrained by its inability to directly observe the effects of introducing the recommendation tool to the market and making it universally accessible. Nevertheless, it can assess the survival rates of Colombian ex-

[15]This specification of PSM does not control by previous product characteristics as this model performs exact matching at the product level, and therefore this dimension is completely absorbed with the PSM specification. However, there is certain degree of heterogeneity within each sector that is not assumed by PSM or any other technique, by construction, as the product is observed at 6-digits HS level. Therefore, PSM assumes certain degree of homogeneity within each sector.

[16]Even after excluding fuel products, which are relevant in value, the products aligning to at least one recommendation still account for 54% of the total export value of products venturing into new markets in 2018, despite representing only 5.1% of the total number of products exported from Colombia to new markets that year.

Table 13: Impact of market recommendation on product growth - PSM

	PSM - Product Export Growth (2018 vs. 2017)
Constant	8.53 **
	(3.14)
Dest. Follow Recomm. (any)	34.73 *
	(14.77)
Observations	162
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*	

Note: This regression uses Propensity Score Matching (PSM) to estimate the impact of coincidentally matching market recommendations, in 2018, on product growth for Colombian exporters. The exact matching is performed at product level (6-digit HS).

porters who matched the recommendation compared to those who did not, providing valuable insights into the efficacy of such a system.

This preliminary evidence may mask potential self-selection bias, as exporters with better market information could potentially exhibit higher levels of productivity, as demonstrated in Görg and Greenaway, 2004, particularly multinational corporations. If this phenomenon indeed explains the empirical findings presented in this paper, advocating for the establishment of a recommendation system accessible to all Colombian exporters would be warranted.

One of the main challenges in this study is that the dataset utilized lacks key firm-level financial variables such as productivity and domestic sales, limiting the ability to fully account for these factors in the analysis. This omission may affect the precision of the results, as firm productivity is often a critical determinant of export entry and survival. Additionally, because the market recommendation system was not followed by exporters, the number of products entering new markets that align with the recommended destinations is relatively small. This scarcity of matched observations presents challenges in generating statistical significance. Future research could address these limitations by incorporating more comprehensive firm-level data, such as productivity measures, employment, and domestic performance. A more detailed dataset would allow for better identification of the mechanisms through which recommendations impact export outcomes. Moreover, expanding the analysis to different contexts or periods where firms might actively adopt a recommendation system could offer richer insights into the tool's effectiveness. These avenues hold promise for deepening the understanding of how tailored recommendations influence exporter dynamics and product expansion after market entry.

3.5 Concluding discussion

This paper uses Colombian customs data to explain the significance of market-level information in elucidating future export survival, complementing the traditional importance given to prior export performance. Therefore, a combination of firm-level information and market-level information is required to understand future successful export decisions. Leveraging highly-accurate Machine Learning models, this paper creates a recommendation tool to simulate an optimal market entry export decision, using complete export market information, for firms trying new destination markets.

The empirical results provide robust evidence of the potential gains from following the model's recommendations. Although firms did not have access to the recommendations at the time, exporters who, by coincidence, entered markets aligned with the model's suggestion saw a 5 percentage point higher probability of survival (34% survival probability) compared to those who selected alternative destinations after previously exporting the same product (29% survival probability). This difference widens to 15 percentage points when comparing these exporters to those entering any new market in the same period (19% survival probability). These findings suggest that adherence to the model's guidance, had it been available, could have substantially improved export survival outcomes. This underscores the value of the recommendation system in enhancing firms' decision-making and increasing the likelihood of export success.

The positive effects stemming from embracing a potential export recommendation tool extend beyond individual exporters, positively impacting overall product growth when at least one exporter in a product category coincidentally follows the model's recommended destination. Employing Ordinary Least Squares (OLS) estimation reveals that products entering markets aligned with at least one model recommendation exhibit a remarkable 102 percentage points increase in growth compared to products that entered non-recommended markets or lacked an associated recommendation. Similarly, a comprehensive Event Study corroborates this result, showing a 35 percentage point increase in product growth when at least one exporter followed the recommended destination. Furthermore, a comparative analysis using Propensity Score Matching (PSM) supports these findings, indicating that products exported to recommended markets consistently experienced significantly higher growth than those entering non-recommended markets, and its size is in line with the result obtained in the Event Study. These robust evidence, obtained through multiple analytical approaches, highlight that products adhering to destination recommendations experience significantly higher growth compared to growth of same products at destinations without such guidance or recommendations. This evidence underscore the substantial and consistent positive impact of heeding market recommendations on product growth within the Colombian export landscape, had exporters followed the guidance provided by the model.

These results provide preliminary evidence that exporters engage in temporary trade as a form of experimentation and to resolve knowledge uncertainty and incomplete information when selecting destinations for market incursions, aligning with findings from previous studies such as Békés and Muraközy, 2012 and Esteve-Pérez, 2021. This perspective aligns with the arguments put forth by Hausmann and Rodrik, 2003, who contend that deficiencies in the discovery phase among developing countries serve as a significant factor contributing to limited export success. Furthermore, my findings build upon the insights of Besedeš and Prusa, 2011, who propose that informational uncertainty may elucidate why numerous export entries terminate prematurely.

Finally, policies aimed at enhancing exporters' ability to optimally identify destination markets by leveraging the collective experiences of diverse firms could significantly benefit Colombian economy. In contexts where firms often lack access to high-quality market information, strategies such as an export recommendation tool, combined with support for off-shore activities like trade fairs and information-sharing initiatives, could drive not only higher survival rates but also broader product growth at the national level. Given its low implementation cost, the recommendation tool could be made available by both private and public entities, helping firms overcome informational barriers that often impede export diversification in developing countries, which can further reduce fixed costs associated with market entry. These initiatives have the potential to equip exporters with more effective decision-making tools, addressing persistent market frictions and contributing to both firm-level success and collective growth, thereby advancing export-led development strategies critical to long-term economic progress.

Conclusion

This thesis presents a comprehensive analysis of the impact of significant events and policies on economic and social outcomes, employing quasi-natural experiments to derive insights in near real-time. The research spans two key areas: international trade and policy evaluation, analyzed through the lenses of the Brexit referendum, the COVID-19 pandemic, and the development of an export recommendation system, each contributing unique perspectives on how shocks and policies influence economic dynamics.

One of the main contributions of this thesis is the demonstration of anticipatory effects of political decisions, such as the Brexit referendum, on a nation's economy. The research highlights how significant economic repercussions can occur even before the formal implementation of a policy, stressing the importance for policymakers to consider these early impacts when planning and communicating major decisions.

Additionally, this thesis introduces an innovative methodology for estimating individual treatment effects in the context of economy-wide shocks, specifically focusing on export dynamics during crises like the COVID-19 pandemic. By applying Machine Learning techniques to construct counterfactual scenarios, the research provides a detailed understanding of how such shocks impact different exporters, offering crucial insights for designing targeted policies that enhance resilience and support vulnerable sectors.

Another significant contribution is the identification of important factors predicting exporter survival, and the development and evaluation of an export recommendation system tailored to the firm-product-market level. This tool not only improves export survival rates but also contributes to national product growth, providing a practical solution for overcoming market entry barriers. The system demonstrates the potential of data-driven approaches to support sustainable economic development, aligning with broader policy goals such as poverty reduction and growth promotion.

Overall, this thesis contributes to the field of international economics by offering new insights into how major economic events and policies affect trade and economic dynamics. It underscores the importance of anticipatory effects, the heterogeneity of impacts across different sectors, and the value of targeted, data-driven interventions. The findings have important implications for policymakers, suggesting that timely and informed decisions can mitigate adverse effects and promote resilience in the face of economic shocks. Through these contributions, the research offers valuable tools and frameworks for both academic inquiry and practical policy application, with the potential to inform future efforts to navigate economic challenges and foster sustainable development.

While this thesis offers valuable insights into the impacts of significant events and policies on economic dynamics, it is not without limitations. One clear limitation is the reliance on quasi-natural experiments, which, while powerful, may not fully capture the complexity and multifaceted nature of the events studied. For instance, the Synthetic Control Method and Machine Learning techniques used to construct counterfactuals are subject to assumptions and model selection biases that could influence the robustness of the findings. Additionally, the generalizability of the results may be constrained by the specific contexts of the UK, Colombia, and the particular datasets used. The export recommendation system, while innovative, is tailored to firm-product-market levels in a specific country, and its applicability to different markets or sectors might require further refinement. Lastly, the dynamic nature of global trade and political environments means that the long-term effects of the policies analyzed may evolve, potentially necessitating ongoing analysis to validate or update the conclusions drawn here.

Appendix A

Appendix to Chapter 1

This chapter is based on the work 'The economic Cost of a Referendum. The case of Brexit' in collaboration with Francesco Serti (see Ortiz-Giménez and Serti (2020)).

A.1 Data and sources

Table 14 describes the data and the sources which we used to perform the study. The database is composed by quarterly data and yearly data predictors. To create a quarterly database we turned the yearly data into quarterly ones. To do this, we copied the yearly data value and introduced it for the four quarters that form a year. With this treatment of the data we lose some variability in the trend of the predictors, but we do not change the real value of the data. The predictors that were originally obtained in yearly frequency are: DNPOP (Density of population), SCSEC (School enrollment, secondary) and the data measuring Foreign Direct Investment (FDII and FDIO). However, our main predictor, real GDP as well as BCI (Business Confidence Index) were originally obtained in quarterly frequency.

Table 14: Data and sources

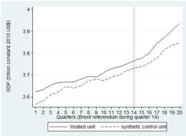
Variable	Description	Source
GDP	Real GDP (US$ millions 2010),by expenditure approach.	OECD
DNPOP	Population density (people per sq. km of land area).	World Bank
SCSEC	School enrollment, secondary (% gross).	"World Development Indicators" from the World Bank
FDII	Foreign Direct Investment inflows (% of GDP)	OECD
FDIO	Foreign Direct Investment outflows (% of GDP)	OECD
BCI	Business Confidence Index	OECD

A.2 European Spillovers

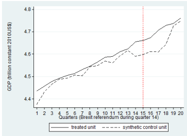
We first try to estimate the synthetic counterfactual for each European country using the baseline donor pool of countries (i.e. without European countries in the donor pool). The results of this exercise are shown in Figure 23. For some of these countries the SCM is not able to reconstruct the GDP pre-treatment path (i.e. the optimization routine is not able to find the minimum of equation 1.1). Among the rest of European countries to which we have successfully applied the SCM, the effect of the treatment could be considered significant only for those countries that present good pre-treatment match, and diverging series after the treatment period. Therefore, there are only four countries that could potentially suffer from spillover effects (indirect effects from Brexit). These countries are Netherlands, Norway, Poland and Switzerland. To evaluate the significance of these effects we do exact inference using the MSPE ratio, as in the main text.

Figure 24 shows that for none of these countries the effect of Brexit can be considered significant. The probability of choosing a country at random with a MSPE ratio as high as the one of Netherlands (Norway, Poland or Switzerland) is $2/11 \simeq 0.18$, that is above the conventional 10% level of significance used in statistics. Then, we cannot reject the null hypothesis and cannot conclude that Brexit referendum have had a significant effect on these European economies.

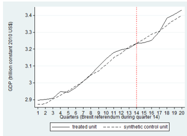
Figure 23: European Spillover Effects.



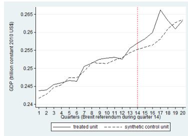
(a) Austria



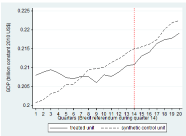
(b) Belgium



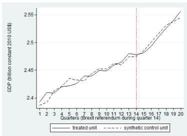
(c) Czech Republic



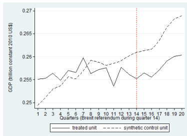
(d) Denmark



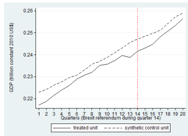
(e) Finland



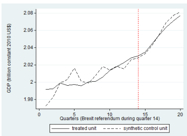
(f) France



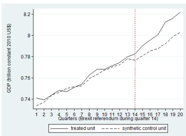
(g) Greece



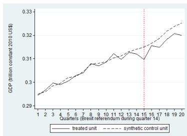
(h) Hungary



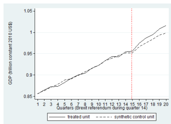
(i) Italy



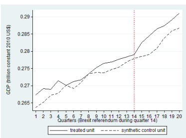
(j) Netherlands



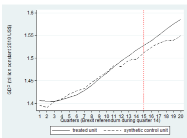
(k) Norway



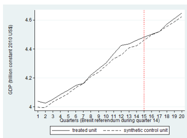
(l) Poland



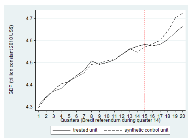
(m) Portugal



(n) Spain



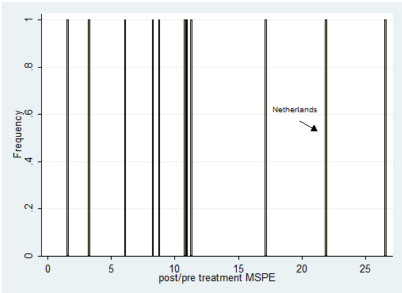
(o) Sweden



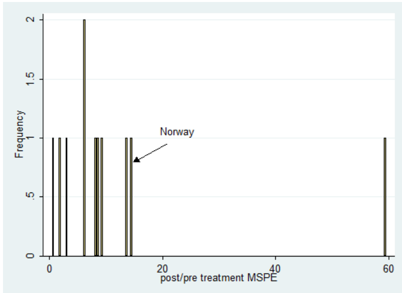
(p) Switzerland

Note: In this figure we assign the treatment to the European countries (that are not the UK) and we estimate its counterfactual using the baseline donor pool (with only non-European countries). This exercise shows that there are no clear individual spillover effects on the European countries analyzed. However there are four countries that present good pre-treatment matching and a jump on their GDP series and therefore they could potentially suffer from indirect effects of Brexit referendum. These countries are Netherlands, Norway, Poland and Switzerland and we will analyze them in depth in Figure 24.

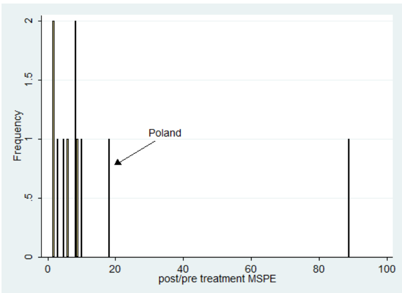
Figure 24: MSPE in European countries.



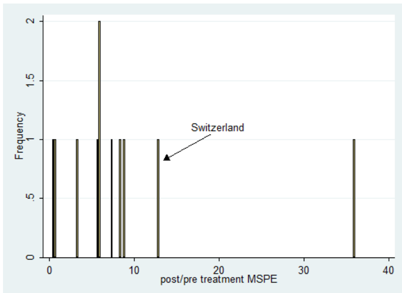
(a) Netherlands



(b) Norway



(c) Poland



(d) Switzerland

Note: This figure shows that the results of in space placebo studies in terms of MSPE ratios for selected indirectly treated European countries.

Appendix B

Appendix to Chapter 2

This chapter is based on the work 'Assessing the Impact of COVID-19 on Trade: a Machine Learning Counterfactual Analysis' in collaboration with Marco Dueñas, Massimo Riccaboni and Francesco Serti (see Dueñas et al. (2021)).

B.1 The Colombian economy amidst the COVID-19 crisis

Colombia exports little compared to other countries in Latin America with similar development levels. In recent years, the share of total exports of Colombian GDP has oscillated around 15%, well below other countries in the region such as Chile and Mexico (Cepeda-López et al., 2019). Colombia started to open its economy in the 1990s with several market-oriented reforms to liberalize financial and capital markets. Although the Colombian economy was still relatively closed during most of the twentieth century (Ocampo and Tovar, 2000), it has been strongly affected by the global financial crisis in 2008-2009 (Zuluaga et al., 2009). Nowadays, despite the growing number of trade agreements, partners, and volume of trade, the integration of Colombia into world trade markets is still modest.

The main reason behind Colombia's poor performance is the low diversification of trade, with a prevalence of primary products, because of the relative abundance of natural resources and low-skilled labor. Besides, the emergence of raw products derived from mining has gained a larger share in total exports, reducing the importance of other products such as coffee, bananas, flowers, some labor-intensive manufactures, and petrochemicals (Garavito et al., 2020).

Since the outbreak of the COVID-19 pandemic, Colombia implemented early measures to contain the spread of COVID-19 and prepare the health system and mitigate the economic and social impact. The Colombian government issued non-compulsory requests for remote working to private companies on February 24, 2020; schools and universities were closed on March 16. On March 25, when there were less than a dozen deaths, the government implemented a complete and mandatory lockdown until April 13. During this period, only a few essential activities – such as health services, public services, communications, banking and financial services, food production, pharmaceuticals, and cleaning and disinfection products – were excluded. The partial lockdown implementation–between April 27 and May 11–allowed a gradual restoration of mobility, enabling a set of non-essential activities under security guidelines and protocols to guarantee social distancing. Most manufacturing activities were gradually allowed at this stage, while non-authorized activities were restricted to market their products through electronic commerce platforms. Finally, from May 28, restrictions to the services sector have been lifted, and on September 1, the government announced the confinement end, and airports were opened.

To better cope with the emergency, Colombian authorities have introduced transitory provisions to secure international trade of essential products. Along with the lockdown measures, medicines, supplies, and equipment in the health sector had zero-tariff for six months. Besides, the export and re-export of these products were forbidden. There was a zero-tariff from April 7 to June 30 for raw materials such as maize, sorghum, soybeans, and soybean cake.

B.2 Data

Table 15: Predictors for exporters success

Variable	Description	Source
	Models: SUM and SAM	
NP, ND, NO	Number of products exported by, number of destinations where a company exports, and number of import origin countries for an exporter in a given month, respectively.	Authors' own elaboration from Colombian Customs Office (DIAN).
HH_p, HH_d	Product-Herfindahl Index, and Destination-Herfindahl Index. Measure the concentration of products at 6-digits HS, and the concentration at destination by company-month, respectively.	Authors' own elaboration from Colombian Customs Office (DIAN).
Total value (exports)	Free on board value of the export transaction in US dollars for each company-month.	Colombian Customs Office (DIAN)
Total value (imports)	Free on board value of the import transaction in US dollars for each company-month.	Colombian Customs Office (DIAN)
Size	4 class dummies classifying firms according to the quartiles of the firm-level (Q1, Q2, Q3 and Q4) distribution of the total monthly value of exports (in $\ln$).	Authors' own elaboration
Destination	Factor variable with one level (dummy variable) for each destination country where Colombian exporters operate by month.	Colombian Customs Office (DIAN)
Origin	Factor variable with one level (dummy variable) for each import origin country where Colombian exporters operate by month.	Colombian Customs Office (DIAN)
Continent	Factor variable with one level (dummy variable) for each continent where Colombian exporters operate.	Authors' own elaboration
Department	Factor variable with one level (dummy variable) for each department (region) in Colombia from which companies operate.	Colombian Customs Office (DIAN)
Means of Transportation	4 class dummies indicating the means of transportation a company use to perform a transaction (land, sea, air, others).	Colombian Customs Office (DIAN)
Sector	99 class dummies classifying company products at 2-digit HS code (corresponding to a HS-chapter).	Authors' own elaboration
Industry	22 class dummies indicating the industries (HS-sections) where companies operate.	Authors' own elaboration from Colombian Customs Office (DIAN).
Sector Experience	Factor variable with one level (dummy variable) for each sector. Takes value 1 in all periods after a company exports for first time in a given sector (reflecting past experience in a sector).	Authors' own elaboration from Colombian Customs Office (DIAN).
Destination Experience	Factor variable with one level (dummy variable) for each destination. Takes value 1 in all periods after a company exports for first time in a given destination (reflecting past experience in a destination).	Authors' own elaboration from Colombian Customs Office (DIAN).
Exporter (importer) Experience	Variable counting the accumulated value exported (imported) in the last twelve months.	Authors' own elaboration from Colombian Customs Office (DIAN).
	Models: SAM (COVID-19 variables)	
Containment Economic Index	We consider the Economic Index from **hale2020variation** that provides a measure of the strength of the economic policies set in place to deal with the pandemic (such as income support and debt relief) for each country in the world. It ranges from 0 to 100. At the firm level we define two variables, one at the export and one at the import side, by taking a weighted average of these country level scores according to the monthly value of exports(imports) that a company ships(source) in every country.[1]	**hale2020variation** and Colombian Customs Office (DIAN).
Containment Government Index	We consider the Government Index from **hale2020variation** that measures the strictness of 'lockdown' style policies that primarily restrict people's behavior. It ranges from 0 to 100. At the firm level we define two variables, one at the export and one at the import side, by taking a weighted average of these country level scores according to the monthly value of exports(imports) that a company ships(source) in every country.	**hale2020variation** and Colombian Customs Office (DIAN).
Containment Health Index	We consider the Health Index from **hale2020variation** that combines 'lockdown' restrictions and closures with measures such as testing policy and contact tracing, short term investment in healthcare, as well investments in vaccine). Ranges from 0 to 100. At the firm level we define two variables, one at the export and one at the import side, by taking a weighted average of these country level scores according to the monthly value of exports(imports) that a company ships(source) in every country.	**hale2020variation** and Colombian Customs Office (DIAN).
Containment Stringency Index	We consider the Stringency Index from **hale2020variation** that records how the response of governments has varied over all indicators, becoming stronger or weaker over the course of the outbreak. Ranges from 0 to 100. At the firm level we define two variables, one at the export and one at the import side, by taking a weighted average of these country level scores according to the monthly value of exports(imports) that a company ships(source) in every country.	**hale2020variation** and Colombian Customs Office (DIAN).
	Models: SUM and SAM (variables only for Logit, Logit-LASSO, and Logit-Ridge)	
Size*Industry	Factor variables with 5 levels for each industry. Takes value 1 when the company size is Q1, value 2 when company size is Q2, value 3 when the size is Q3 and 4 when the size is Q4 while operating in a given industry. However, it takes value 0 if a company is not operating in this industry (for any size level).	Authors' own elaboration from Colombian Customs Office (DIAN).
Size*Sector	Factor variables with 5 levels for each sector. Takes value 1 when the company size is Q1, value 2 when company size is Q2, value 3 when the size is Q3 and value 4 when the size is Q4 while operating in a given sector. However, it takes value 0 if a company is not operating in this sector (for any size level).	Authors' own elaboration from Colombian Customs Office (DIAN).
Size*Means of Transportation	Factor variables with 5 levels for each sector. Takes value 1 when the company size is Q1, value 2 when company size is Q2, value 3 when the size is Q3 and value 4 when the size is Q4 while operating using a given means of transportation. However, it takes value 0 if a company is not operating using this means of transportation (for any size level).	Authors' own elaboration from Colombian Customs Office (DIAN).
Size*Destination	Factor variables with 5 levels for each sector. Takes value 1 when the company size is Q1, value 2 when company size is Q2, value 3 when the size is Q3 and value 4 when the size is Q4 while operating in a given destination. However, it takes value 0 if a company is not operating in this destination (for any size level).	Authors' own elaboration from Colombian Customs Office (DIAN).

* `https://data.europa.eu/euodp/en/data/dataset/covid-19-coronavirus-data`

[1] When an exporter does not import we impute the corresponding internal Index (Economic, Government, Health, and Stringency) of Colombia to create the corresponding import side Index.

Table 16: Sector-Industry mapping

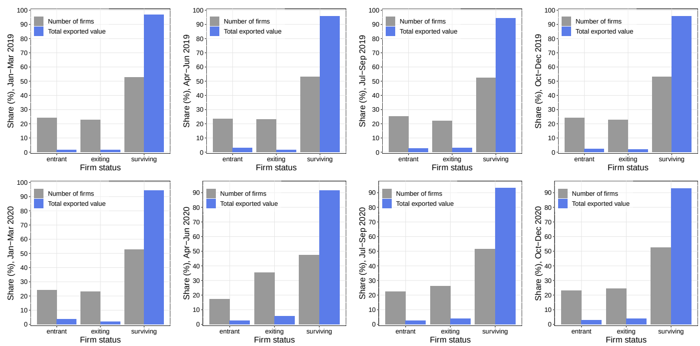
Section (Industry)	Industry Name	HS-Chapter (Sector)
1	Live Animals/ Animal Products	1-5
2	Vegetable Products	6-14
3	Animal or Vegetable Fats/Oils	15
4	Prepared Foodstuffs	16-24
5	Mineral Products	25-27
6	Products of Chemical Industries	28-38
7	Plastics, Rubber	39-40
8	Raw Hides, Skins and Leather	41-43
9	Wood	44-46
10	Paper	47-49
11	Textile	50-63
12	Footwear	64-67
13	Art. of Stone, Cement	68-70
14	Jewelries	71
15	Base Metals	72-83
16	Machinery Equipment	84-85
17	Vehicles	86-89
18	Precision Instruments	90-92
19	Arms	93
20	Miscellaneous Manufactured Articles	94-96
21	Works of Art	97
22	Special Classification Provisions	98-99

Source: Author's elaboration using Pierce and Schott, 2012 tables.

B.3 Additional Descriptive Statistics

Figure 25 shows, separately for the first and second quarter of a year, the percentage of firms that survive, enter or exit the export market and their corresponding shares of total exports. Thus, for a given quarter in 2019 and the corresponding quarter in 2020, we label each firm as *exiting* when it is present in 2019 and absent in 2020, *entrant* when it is absent in 2019 and present in 2020, and *surviving* when it is present in both years. We average the total value exported by each firm during the same quarter of two different years. Then, we sum the individual average value exported according to the firms' status. It turns out that surviving firms play an essential role in explaining total exports: they are around half of the total number of firms in both quarters and account for about 90% of the total export value. The volume lost, during the second quarter of 2020, due to exiting firms is around 5% (assuming they would have exported in 2020 similar export volumes as observed in 2019). Entrant firms almost made up this 5% loss. Despite this, the firms' composition that participates in exports is very different. The number of existing firms in the second quarter of 2020 is much higher than the share of the first quarter of 2020 and the share of 2019 in the same period of the year.

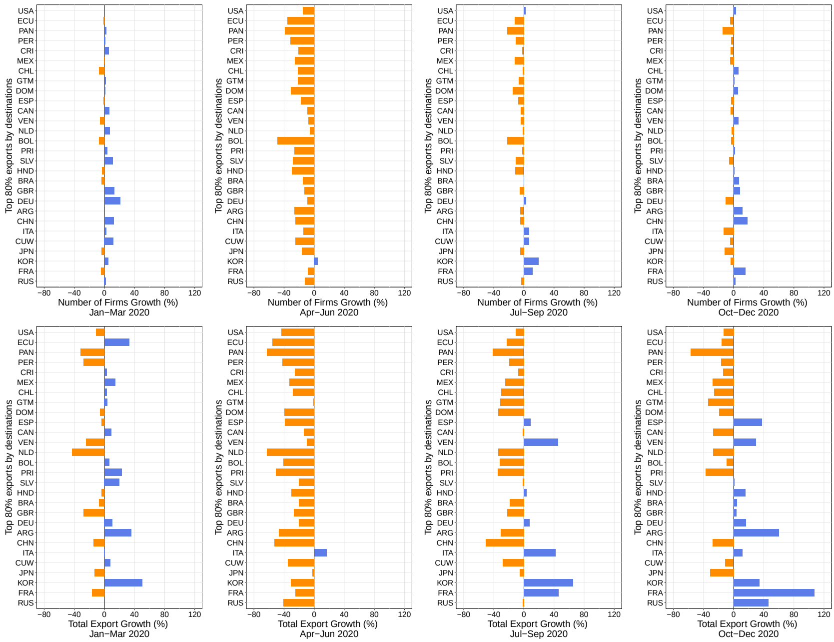
Figure 25: Exporters dynamics in Colombia



Note: Entry-exit dynamics of exporters and total export value by firms that drop, enter or stay active, in 2019 (upper part of the figure) and in 2020 (bottom part of the figure) by quarters. Firm status is defined by looking at the previous year.

Figures 26 and 27 show the growth of the total number of exporters and the growth of the total volume of exports between 2019 and 2020, by country of destination and product sector. We consider the first and the second quarter separately, and we select destinations and product sectors that account for 80% of the total exporters in 2019. In both figures, the product sectors and the destinations are arranged by importance from top to bottom.

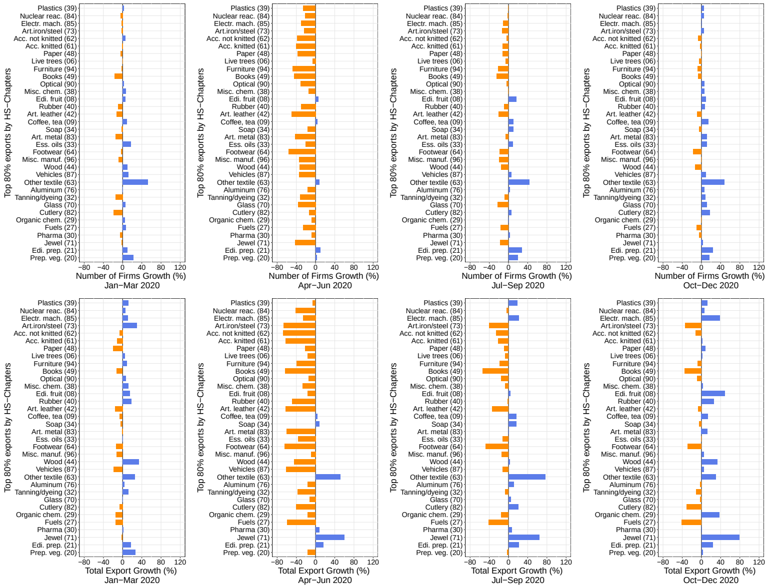
Figure 26: Total number of exporters and value growth by market in 2020



Note: The growth of the total number of exporters and the total value of exports by the destination country for the four quarters of 2020. Orange bars represent negative growth and blue bars positive growth. Destination countries are sorted from top to bottom accordingly to their importance in the share of the number of exporters in 2019.

Figure 26 shows that, compared to the first quarter of 2020, the second quarter of the year is characterized by a severe and pervasive drop in the number of exporting firms and the volume of exports in almost all the destinations reported. Figure 28 shows that the same drop is not observed during the second quarter of 2019. During the third and fourth quarters of 2020, the value exported experienced more volatility than the

Figure 27: Total number of exporters and value growth by sector in 2020



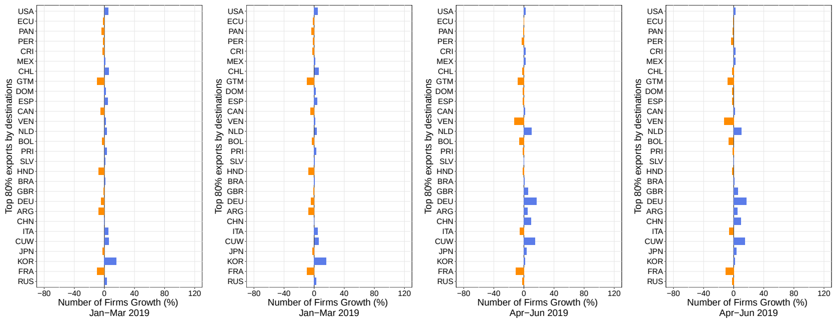
Note: The growth of the total number of exporters and the total value of exports by sector for the four quarters of 2020. Orange bars represent export reductions and blue bars positive export growth. Product sectors are sorted from top to bottom according to their importance in the share of the number of exporters in 2019. Product sectors correspond to the chapters of the HS code in parenthesis and the full name of the chapters is shortened to improve readability.

number of firms. Nevertheless, the latter did not recover to the growth rates of the first quarter of the year.

Exports by product sectors in the second quarter of 2020 (see Figure 27) reveal a generalized decrease in the number of exporting firms and trade values, while the first quarter exhibits very heterogeneous patterns. The sectors that appear to be more severely affected in the second quarter are Footwear (HS64), Leather Articles (HS42), Furniture (HS94), Books (HS49), Articles of Metal (HS83), Knitted and Not-Knitted Accessories (HS61-62), Vehicles (HS87) and Articles of Iron or Steel (HS73). Interestingly, these sectors are relatively more labor-intensive in Colombia, and therefore they could be susceptible to disruptions connected to social distancing. Finally, only for Coffee and Tea (HS08), Other textiles

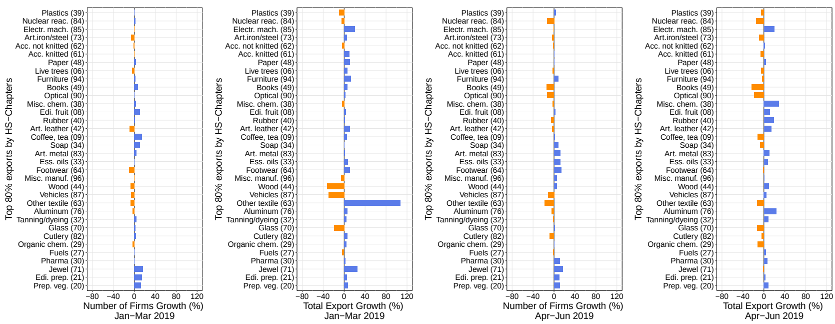
(HS63), and Jewelries (HS71) exports in value significantly grew in the second quarter. Instead, in terms of the number of exporting firms, no product sectors exhibit notable positive dynamics. During the third and fourth quarters of 2020, there is a rapid back to normality in both the growth of value exported and in the number of exporters' growth rate by sector. Figure 29 shows the same figures for 2019, suggesting that in periods without relevant shocks – such as the ones of the first quarter of 2020 – the changes in exports are also very heterogeneous, but there are not such extreme changes.

Figure 28: Total number of exporters and value growth by destination in 2019



Note: The growth of the total number of exporters and the total value of exports by destination country for the first and the second quarters of 2019. Orange bars represent negative growth and blue bars positive growth. Destination countries are sorted from top to bottom accordingly with their importance in the share of number of exporters in 2019.

Figure 29: Total number of exporters and value growth by sector in 2019



Note: The growth of the total number of exporters and the total value of exports by sector for the first and the second quarters of 2019. Orange bars represent export reductions and blue bars positive export growth. Product sectors are sorted from top to bottom accordingly with their importance in the share of number of exporters in 2019. Product sectors correspond to the chapters of the HS-code in parenthesis, the full name of the chapters is shortened to improve readability.

Appendix C

Appendix to Chapter 3

This chapter is based on the work 'Exports' Survival in New Markets: A firm-level export recommendation model'. In particular, this chapter is a solo-author paper that provides value to Colombian exporters, presenting a market recommendation system, and helps the literature to better understand the impact that export recommendation can generate on export survival and product growth.

C.1 Data

Table 17: Predictors for exporters success

Variable	Description	Source
	Models: SUM and SAM	
NP, ND, NO, NV	Number of products exported, number of destinations where a firm exports, number of import origin countries for an exporter in a given month, and number of product-market (varieties) combinations respectively.	Authors' own elaboration from Colombian Customs Office (DIAN).
Total value (exports)	Free on board value of the export transaction in US dollars for each firm-product-destination-year.	Colombian Customs Office (DIAN)
Size	4 class dummies classifying firms according to the quartiles of the firm-level (Q1, Q2, Q3 and Q4) distribution of the total monthly value of exports (in $\ln$).	Authors' own elaboration
Destination	Factor variable with one level (dummy variable) for each destination country where Colombian exporters operate by month.	Colombian Customs Office (DIAN)
Origin	Factor variable with one level (dummy variable) for each import origin country where Colombian exporters operate by month.	Colombian Customs Office (DIAN)
Department	Factor variable with one level (dummy variable) for each department (region) in Colombia from which firms operate.	Colombian Customs Office (DIAN)
Means of Transportation	4 class dummies indicating the means of transportation an exporter use to perform a transaction (land, sea, air, others).	Colombian Customs Office (DIAN)
Sector	99 class dummies classifying exporter products at 2-digit HS code (corresponding to a HS-chapter).	Authors' own elaboration
Industry	22 class dummies indicating the industries (HS-sections) where firms operate.	Authors' own elaboration from Colombian Customs Office (DIAN).
Destination Experience	Variables (one per potential destination) indicating the cumulative exported value to a specific destination, represented as a percentage of the total previous exports to all destinations up to one year before the data observation.	Authors' own elaboration from Colombian Customs Office (DIAN).
Experience (SP-OM)	Variable representing the duration, in years, in which a firm has been exporting the identical product to other destinations up to one year prior to the data observation.	Authors' own elaboration from Colombian Customs Office (DIAN).
Experience (OP-OM)	Variable representing the duration, in years, in which a firm has been exporting different product to various destinations up to one year prior to the data observation.	Authors' own elaboration from Colombian Customs Office (DIAN).
Old exporter	Dummy variable that assumes a value of 1 if the exporter had engaged in exporting at any given point in time; otherwise, it assumes a value of 0.	Authors' own elaboration from Colombian Customs Office (DIAN).
Market Share	Market share of firm i in a product-destination-year as a percentage of the total exports in year t.	Authors' own elaboration from Colombian Customs Office (DIAN).
Product-Market Portfolio	Percentage that represents a product-destination-year of firm's portfolio in year t.	Authors' own elaboration from Colombian Customs Office (DIAN).
FP	Variable that takes value 1 if it is the first time that firm i exports a given product.	Authors' own elaboration from Colombian Customs Office (DIAN).
N. firms in P, M, or PM in $t-1$	Three different variables counting the number of firms (different than the one observed) in period $t-1$ that exported the same product (P), market (M), or product-market (PM), respectively, than the observed exporter.	Authors' own elaboration from Colombian Customs Office (DIAN).
N. firms in P, M, or PM in t	Three different variables counting the number of firms (different than the one observed) in period t that exported the same product (P), market (M), or product-market (PM), respectively, than the observed exporter.	Authors' own elaboration from Colombian Customs Office (DIAN).
Acc. value in P	Variable counting the accumulated exported firm's value in product (P) until period $t-1$.	Authors' own elaboration from Colombian Customs Office (DIAN).
Experienced exporter	Dummy variable that takes value 1 when the exporter i has been observed in the dataset before period t.	Authors' own elaboration from Colombian Customs Office (DIAN).
Experience exporting P	Variable counting the total number of years that firm i has been exporting product (P) until period $t-1$.	Authors' own elaboration from Colombian Customs Office (DIAN).
N. firms exporting in the same P, M, PM	Number of firms exporting the same product (P), in the same market (M), or in the same product-market (PM), respectively, in year t.	Authors' own elaboration from Colombian Customs Office (DIAN).
N. of P in the same M	Number of products exporting to same market (M), in year t.	Authors' own elaboration from Colombian Customs Office (DIAN).
N. of M with the same P	Number of markets exporting the same product (P), in year t.	Authors' own elaboration from Colombian Customs Office (DIAN).
Exporter (importer) Experience	Variable counting the accumulated value exported (imported) in the last twelve months.	Authors' own elaboration from Colombian Customs Office (DIAN).

* https://data.europa.eu/euodp/en/data/dataset/covid-19-coronavirus-data

[1] When an exporter does not import, I impute the corresponding internal Index (Economic, Government, Health, and Stringency) of Colombia to create the corresponding import side Index.

C.2 Variable Importance

See Figure 30.

C.3 Robustness: Information Removal

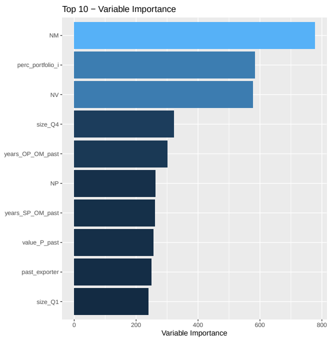
In this exercise, I conduct four Random Forest (RF) regressions as outlined in Table 10 following a random removal of 10% of the variables, repeated 150 times. This iterative process serves to reinforce the inherent robustness of the RF methodology, which guards against overfitting by repeatedly generating diverse trees with distinct partitioning schemes. The upper part of Table 18 illustrates the average proportion of market-related variables among the Top-10 most influential factors. This proportion reaches its peak when the model predicts survival within the product-market dimension, with nearly 60%, on average, of the Top-10 most influential variables embodying market-knowledge information. Conversely, the bottom part shows the average Goodness of Fit metrics, accuracy, sensitivity, and specificity for each model, respectively. These metrics also survive the iterative removal of 10% variables by keeping similar levels to the ones achieved in Table 10.

Table 18: Robustness Check - Iterative removal of variables

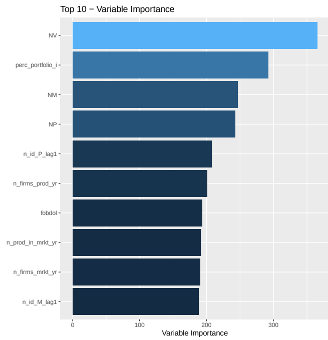
Continue exporting, after market entry, in period t+2 with the same...				
	id	**Product**	**Market**	**PM**
% Market-Knowledge vars. among Top-10 (avg.)	0.4%	15.27%	49.4%	59.1%
Accuracy (avg.)	78%	79%	70%	74%
Sensitivity (avg.)	79%	82%	73%	77%
Specificity (avg.)	78%	75%	62%	59%

Note: Iterative removal of 10% variables 150 times: Average proportion of market-related variables among the Top-10 most relevant factors and Goodness of Fit measures for export survival at Firm, Firm-Product, Firm-Market, or Firm-Product-Market level, respectively, two periods after a market entry. Prediction (2017), train (2010-2016).

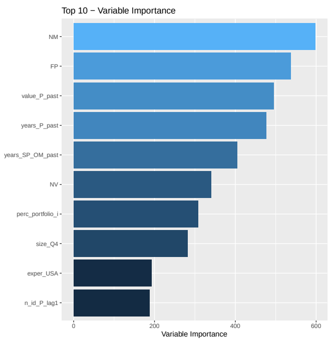
Figure 30: Variable Importance - Gini Index (Impurity)



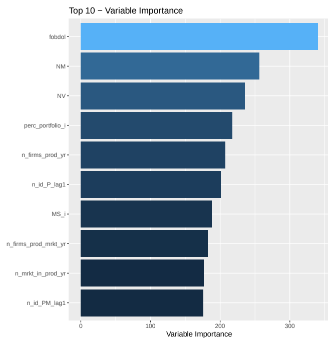
Survival as exporter at period $t+2$ after trying a market incursion (M) in t.



Survival in (M) at period $t+2$ after trying a market incursion (M) in t.



Survival in (P) at period $t+2$ after trying a market incursion (M) in t.



Survival in (PM) at period $t+2$ after trying a market incursion (M) in t.

Note: Variable Importance ranking measured as the Impurity (reduction) from the Gini Index. Ranking obtained after predicting survival rates of exporters in period $t+2$, following different definitions, after a market incursion in t.

Bibliography

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program". In: *Journal of the American statistical Association* 105.490, pp. 493–505.

— (2015). "Comparative politics and the synthetic control method". In: *American Journal of Political Science* 59.2, pp. 495–510.

Abadie, Alberto and Javier Gardeazabal (2003). "The economic costs of conflict: A case study of the Basque Country". In: *American economic review* 93.1, pp. 113–132.

Albornoz, Facundo, Sebastián Fanelli, and Juan Carlos Hallak (2016). "Survival in export markets". In: *Journal of International Economics* 102, pp. 262–281.

Albornoz, Facundo et al. (2023). "Sequentially exporting products across countries". In: *Journal of International Economics* 142, p. 103735.

Alfaro-Urena, Alonso et al. (2023). *Firm export dynamics in interdependent markets*. Tech. rep. National Bureau of Economic Research.

Alvarez, Roberto, Hasan Faruq, and Ricardo A. López (2013). "Is previous export experience important for new exports?" In: *The Journal of Development Studies* 49.3, pp. 426–441.

Antràs, Pol and Davin Chor (2022). "Global Value Chains". In: *Handbook of International Economics*. Vol. 5. Elsevier.

Antràs, Pol, Stephen J. Redding, and Esteban Rossi-Hansberg (2023). "Globalization and pandemics". In: *American Economic Review* 113.4, pp. 939–981.

Atkin, David, Amit K Khandelwal, and Adam Osman (2017). "Exporting and firm performance: Evidence from a randomized experiment". In: *The quarterly journal of economics* 132.2, pp. 551–615.

Baldwin, Richard and Eiichi Tomiura (2020). “Thinking ahead about the trade impact of COVID-19”. In: *Economics in the Time of COVID-19* 59.

Békés, Gábor and Balázs Muraközy (2012). “Temporary trade and heterogeneous firms”. In: *Journal of International Economics* 87.2, pp. 232–246.

Bernard, Andrew B. and J. Bradford Jensen (2004). “Why some firms export”. In: *Review of economics and Statistics* 86.2, pp. 561–569.

Berthou, Antoine and Sebastian Stumpner (2022). *Trade under lockdown*. Tech. rep. Banque de France Working Paper.

Besedeš, Tibor and Thomas J. Prusa (2006a). “Ins, outs, and the duration of trade”. In: *Canadian Journal of Economics/Revue canadienne d'économique* 39.1, pp. 266–295.

Besedeš, Tibor and Thomas J Prusa (2006b). “Product differentiation and duration of US import trade”. In: *Journal of international Economics* 70.2, pp. 339–358.

Besedeš, Tibor and Thomas J. Prusa (2011). “The role of extensive and intensive margins and export growth”. In: *Journal of development economics* 96.2, pp. 371–379.

Bonadio, Barthélémy et al. (2020). *Global supply chains in the pandemic. National Bureau of Economic Research, No. w27224*.

Born, Benjamin et al. (2019). “The costs of economic nationalism: evidence from the Brexit experiment”. In: *The Economic Journal* 129.623, pp. 2722–2744.

Breiman, Leo (2001). “Random forests”. In: *Machine learning* 45, pp. 5–32.

Breinlich, Holger et al. (2022). *Machine learning in international trade research-evaluating the impact of trade agreements. CEPR Discussion Paper No. DP17325*.

Broocks, Annette and Johannes Van Biesebroeck (2017). “The impact of export promotion on export market entry”. In: *Journal of International Economics* 107, pp. 19–33.

Campos, Nauro F., Fabrizio Coricelli, and Luigi Moretti (2018). “Institutional integration and economic growth in Europe”. In: *Journal of Monetary Economics*.

Card, David and Alan B. Krueger (1993). *Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania*.

Cepeda-López, F. et al. (2019). “Colombian liberalization and integration to world trade markets: Much ado about nothing”. In: *Borradores de Economía; No. 1065*.

Cerqua, Augusto and Marco Letta (2020). "Local economies amidst the COVID-19 crisis in Italy: a tale of diverging trajectories". In: *COVID Economics* 60, pp. 142–171.
Chaney, Thomas (2014). "The network structure of international trade". In: *American Economic Review* 104.11, pp. 3600–3634.
Che, Natasha Xingyuan (2020). *Intelligent export diversification: an export recommendation system with machine learning. IMF Working Paper*.
Chernozhukov, Victor, Iván Fernández-Val, and Ye Luo (2018). "The sorted effects method: discovering heterogeneous effects beyond their averages". In: *Econometrica* 86.6, pp. 1911–1938.
Chernozhukov, Victor et al. (2020). *Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India*. Tech. rep. National Bureau of Economic Research.
Chow, Gregory C (1960). "Tests of equality between sets of coefficients in two linear regressions". In: *Econometrica: Journal of the Econometric Society*, pp. 591–605.
Ciuriak, Dan et al. (2015). *The trade-related impact of a UK Exit from the EU single market. Available at SSRN: https://ssrn.com/abstract=2620718*.
De Lucio, Juan et al. (2022). "Impact of Covid-19 containment measures on trade". In: *International Review of Economics & Finance* 80, pp. 766–778.
Decreux, Yvan and Julia Spies (2016). "Export potential assessments". In: *International Trade Centre*.
Dhingra, Swati et al. (2016). "The impact of Brexit on foreign investment in the UK". In: *BREXIT 2016* 24.
Dhingra, Swati et al. (2017). "The costs and benefits of leaving the EU: trade effects". In: *Economic Policy* 32.92, pp. 651–705.
Doudchenko, Nikolay and Guido W. Imbens (2016). *Balancing, regression, difference-in-differences and synthetic control methods: A synthesis*. Tech. rep. National Bureau of Economic Research.
Dueñas, Marco et al. (2021). "Assessing the impact of covid-19 on trade: a machine learning counterfactual analysis". In: *arXiv preprint arXiv:2104.04570*.
Eckel, Carsten and J. Peter Neary (2010). "Multi-product firms and flexible manufacturing in the global economy". In: *The Review of Economic Studies* 77.1, pp. 188–217.
Espitia, Alvaro et al. (2021). "Pandemic trade: COVID-19, remote work and global value chains". In: *The World Economy*.

Esteve-Pérez, Silviano (2021). "Previous experience, experimentation and export survival: Evidence from firm-product-destination level data". In: *The World Economy* 44.9, pp. 2638–2682.

Esteve-Perez, Silviano, Francisco Requena-Silvente, and Vicente J Pallardo-Lopez (2013). "The duration of firm-destination export relationships: Evidence from Spain, 1997–2006". In: *Economic Inquiry* 51.1, pp. 159–180.

Evenett, Simon (2020). "Sicken thy neighbour: The initial trade policy response to COVID-19". In: *The World Economy* 43.4, pp. 828–839.

Fabra, Natalia, Aitor Lacuesta, and Mateus Souza (2022). "The implicit cost of carbon abatement during the COVID-19 pandemic". In: *European Economic Review* 147, p. 104165. ISSN: 0014-2921. DOI: https://doi.org/10.1016/j.euroecorev.2022.104165. URL: https://www.sciencedirect.com/science/article/pii/S0014292122000939.

Felbermayr, Gabriel and Holger Görg (2020). "Implications of Covid-19 for globalization". In: *KIELER BEITRÄGE ZUR*, p. 3.

Ferman, Bruno and Cristine Pinto (2021). "Synthetic controls with imperfect pretreatment fit". In: *Quantitative Economics* 12.4, pp. 1197–1221.

Ferman, Bruno, Cristine Pinto, and Vitor Possebom (2020). "Cherry picking with synthetic controls". In: *Journal of Policy Analysis and Management* 39.2, pp. 510–532.

Fugazza, Marco and Ana Cristina Molina (2016). "On the determinants of exports survival". In: *Canadian Journal of Development Studies/Revue canadienne d'études du développement* 37.2, pp. 159–177.

Garavito, Aarón et al. (2020). "Ingresos externos corrientes de Colombia: desempeño exportador, avances y retos". In: *Revista Ensayos Sobre Política Económica; No. 95, julio 2020. Pág.: 1-81*.

Görg, Holger and David Greenaway (2004). "Much ado about nothing? Do domestic firms really benefit from foreign direct investment?" In: *The World Bank Research Observer* 19.2, pp. 171–197.

Grossman, Gene M., Elhanan Helpman, and Hugo Lhuillier (2023). "Supply chain resilience: Should policy promote international diversification or reshoring?" In: *Journal of Political Economy* 131.12, pp. 3462–3496.

Hale, Thomas et al. (2021). "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)". In: *Nature human behaviour* 5.4, pp. 529–538.

Halpern, László, Miklós Koren, and Adam Szeidl (2015). "Imported inputs and productivity". In: *American Economic Review* 105.12, pp. 3660–3703.

Haskel, Jonathan, Sonia C. Pereira, and Matthew J Slaughter (2007). "Does inward foreign direct investment boost the productivity of domestic firms?" In: *The review of economics and statistics* 89.3, pp. 482–496.

Hausmann, Ricardo and Dani Rodrik (2003). "Economic development as self-discovery". In: *Journal of development Economics* 72.2, pp. 603–633.

Hess, Wolfgang and Maria Persson (n.d.). "Exploring the duration of EU imports". In: *Review of World Economics* 147 (), pp. 665–692.

Hidalgo, César A. et al. (2007). "The product space conditions the development of nations". In: *Science* 317.5837, pp. 482–487.

Holland, Paul W. (1986). "Statistics and causal inference". In: *Journal of the American statistical Association* 81.396, pp. 945–960.

James, Gareth et al. (2013). *An introduction to statistical learning*. Vol. 112. Springer.

Jun, Bogang et al. (2020). "Bilateral relatedness: knowledge diffusion and the evolution of bilateral trade". In: *Journal of Evolutionary Economics* 30, pp. 247–277.

Kierzenkowski, Rafal et al. (2016). *The economic consequences of brexit*. Tech. rep. OECD Publishing.

Kramarz, Francis, Julien Martin, and Isabelle Mejean (2020). "Volatility in the small and in the large: The lack of diversification in international trade". In: *Journal of international economics* 122, p. 103276.

Lafrogne-Joussier, Raphael, Julien Martin, and Isabelle Mejean (2022). "Supply shocks in supply chains: Evidence from the early lockdown in China". In: *IMF Economic Review*, pp. 1–46.

Latorre, María C. et al. (2019). "Brexit: Everyone loses, but Britain loses the most". In: *Peterson Institute for International Economics Working Paper* 19-5.

Liu, Xuepeng, Emanuel Ornelas, and Huimin Shi (2021). "The Trade Impact of the Covid-19 Pandemic". In: *The World Economy*.

Miller, Douglas L. (2023). "An Introductory Guide to Event Study Models". In: *Journal of Economic Perspectives* 37.2, pp. 203–230.

Montinari, Letizia, Massimo Riccaboni, and Stefano Schiavo (2021). "Innovation, trade and multi-product firms". In: *Canadian Journal of Economics/Revue canadienne d'économique*.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler (2019). "Extended gravity". In: *The Review of Economic Studies* 86.6, pp. 2668–2712.

Munch, Jakob and Georg Schaur (2018). “The effect of export promotion on firm-level performance”. In: *American Economic Journal: Economic Policy* 10.1, pp. 357–87.
Navas, Antonio, Francesco Serti, and Chiara Tomasi (2020). “The role of the gravity forces on firms’ trade”. In: *Canadian Journal of Economics/Revue canadienne d’économique* 53.3, pp. 1059–1097.
Nitsch, Volker (2009). “Die another day: Duration in German import trade”. In: *Review of World Economics* 145, pp. 133–154.
— (2022). “Covid-19 and international trade: Evidence from New Zealand”. In: *Economics Letters*, p. 110627.
Ocampo, José Antonio and Camilo Tovar (2000). “Colombia in the Classical Era of ‘Inward-Looking Development’, 1930–74”. In: *An economic history of twentieth-century Latin America*. Ed. by E Cárdenas, JA Ocampo, and R Thorp. Springer, pp. 239–281.
Olarreaga, Marcelo, Stefan Sperlich, and Virginie Trachsel (2020). “Exploring the Heterogeneous Effects of Export Promotion”. In: *The World Bank Economic Review* 34.2, pp. 332–350.
Ortiz, Gabriela and María C. Latorre (2018). “The economic impact of potential migration policies in the UK after Brexit”. In: *Contemporary Social Science*, pp. 1–18.
Ortiz-Giménez, Víctor and Francesco Serti (2020). “The economic cost of a referendum: The case of Brexit”. In: *Resúmenes de la III Jornada de Investigación en Internacionalización*. Universidad Antonio de Nebrija, pp. 31–37.
Perron, Pierre (1997). “Further evidence on breaking trend functions in macroeconomic variables”. In: *Journal of econometrics* 80.2, pp. 355–385.
Pierce, Justin R. and Peter K Schott (2012). “A concordance between ten-digit US Harmonized System Codes and SIC/NAICS product classes and industries”. In: *Journal of Economic and Social Measurement* 37.1-2, pp. 61–96.
Roberts, Mark J. and James R. Tybout (1997). “The decision to export in Colombia: An empirical model of entry with sunk costs”. In: *The american economic review*, pp. 545–564.
Rosenbaum, Paul R. and Donald B. Rubin (1983). “The central role of the propensity score in observational studies for causal effects”. In: *Biometrika* 70.1, pp. 41–55.
Treasury, HM (2016). “HM Treasury analysis: the immediate economic impact of leaving the EU”. In: *May, Cm* 9292.

Van Beveren, Ilke, Andrew B. Bernard, and Hylke Vandenbussche (2012). *Concording EU trade and production data over time*. Tech. rep. National Bureau of Economic Research.
Van Biesebroeck, Johannes, Jozef Konings, and Christian Volpe Martincus (2016). "Did export promotion help firms weather the crisis?" In: *Economic Policy* 31.88, pp. 653–702.
Vandenbussche, Hylke, William Connell, and Wouter Simons (2022). "Global value chains, trade shocks and jobs: An application to Brexit". In: *The World Economy* 45.8, pp. 2338–2369.
Varian, Hal R. (2016). "Causal inference in economics and marketing". In: *Proceedings of the National Academy of Sciences* 113.27, pp. 7310–7315.
Volpe Martincus, Christian and Jerónimo Carballo (2008). "Is export promotion effective in developing countries? Firm-level evidence on the intensive and the extensive margins of exports". In: *Journal of International Economics* 76.1, pp. 89–106.
— (2010). "Entering new country and product markets: does export promotion help?" In: *Review of World Economics* 146.3, pp. 437–467.
World Bank (2020). *Global Economic Prospects, June 2020. Washington, DC: World Bank. Available at:* `https://openknowledge.worldbank.org/handle/10986/33748`.
Zuluaga, Óscar Iván et al. (2009). "Informe de la Junta Directiva al Congreso de la República – Marzo de 2009". In: Banco de la República de Colombia.