# Assessing the heterogeneous impact of economy-wide shocks. Evidence from Colombian firms during the COVID-19 crisis.

Marco Dueñas, Federico Nutarelli, Víctor Ortiz, Massimo Riccaboni and Francesco Serti

IMT School for Advanced Studies Lucca

**Seminario de Investigación de Facultad, Universidad Complutense de Madrid**

November 24, 2023

# Introduction

- We present a **methodology to study the heterogeneous effects of economy-wide shocks.**
- This methodology is applicable in scenarios where the **pervasive** nature of the shock and the existence of **unobserved interactions** between units hinders the **identification of a control group unaffected** by the shock (impossible to apply a DID strategy), as well as the ex-ante definition of the **intensity of the shock's exposure** of each unit (necessary for a continous DID strategy).
- We apply it to identify **the heterogeneous effect of COVID-19 on a firm's probability of survival in the export markets**.
  - All firms are eventually directly or indirectly exposed to COVID-19
  - Defining ex-ante the intensity with which each firm is exposed to the treatment is challenging

# Contributions

- This paper's first contribution is exploring the **effectiveness of different ML techniques in predicting the probability of Colombian firms to survive** in the export market under two different scenarios: a COVID-19 setting and a non-COVID-19 counterfactual situation.
- This prediction step is the **first stage of our Causal Machine Learning (ML) strategy** to estimate the heterogeneous effect of global shocks.
- **By comparing these estimated potential outcomes we obtain an estimate of the the COVID-19 effect on the probability of survival at the firm level**.
- Then, we study the **heterogeneity** of the COVID-19 effects according to firms' characteristics.

# Contributions

- The **traditional approach** splits the sample into groups to assess the significance of the difference in the treatment effects of the groups.
- Unfortunately, this approach is prone to **overfitting**, finding statistically significant differences out of all possible splits might be entirely due to random noise, and is unable to **capture complex, higher-order interactions** between treatment and baseline characteristics.
- By **adapting recent Causal ML tools (Chernozhukov et al. 2018, 2020) to a setting without a control group**, we use the estimated firm-level effects stemming from our ML counterfactual empirical model
  - to classify firms in two groups, the most and the least affected by the COVID-19 shock,
  - and then we compare their average characteristics.
- This method avoids to avoid overfitting/p-hacking/fishing issues and is robust to possible non-linearities and interactions effects with other variables.

# Data

1. Exporters' characteristics from the **Monthly export transactions data of Colombian Customs Office (DIAN)** for 2018-2020, considering: number of destinations and products exported/imported, total export/import value, means of transportation, sector, location,...
2. To try to measure the exposure to COVID-19 demand and supply shock, we use of four indexes (**Government Response Trackers**, ranging from 0 to 100) representing the strength of the measures taken by countries to contain the COVID-19 outbreak.

    - We build two time-varying variables at the firm level 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.
    - These exposure variables are used just as some possible determinants of treatment effect heterogeneity, they are not imposed ex-ante as the only treatment intensity determinants as it would be done in a continuous DID strategy.

# Methodology I: Potential Outcomes

- The outcome that we want to analyze is whether a firm that was exporting in a given month in 2019 **will export again in the same month of 2020**.
- For each month, we build two different models:
  - **Shock Unaware Machine** (SUM): it is the model that we use to estimate the **potential outcome in case of no treatment**, which does not consider the COVID-19 information (only 2018-2019 data). The term "Machine" refers to the fact that the **counterfactual** has been constructed through ML techniques.
  - **Shock Aware Machine** (SAM): it is the model for the **potential outcome in case of treatment**: it is fully aware of all the available information related to the COVID-19 scenario (i.e., firms behaviour in 2020 and measures at economic, health and government level, summarized in the different Indexes) This **summarizes the information on the observed COVID-19 scenario** and expresses it in a metric that is comparable with the SUM.
  (**both are probabilities**) Emp.Strat.

# Methodology II: Potential Outcomes

- We **train SUM** by using the characteristics of exporters observed in 2018 to explain their export behavior in 2019.
- We choose the best performing predictive SUM model (out of sample) using **cross-validation techniques (i.e., K-fold method)**.
- **We apply the "best performing" SUM (trained using data in 2018-2019) to predict the 2020 outcome** for firms exporting in 2019 (the counterfactual).
- The main assumption is that **we can learn what would have happened in 2020 without the COVID-19 by exploiting the observed firm behavior and characteristics in 2018-2019.**
- The SAM machine instead considers the exporters operating in the market in 2019 and use their observed dynamics in 2020: it is just a probabilistic picture of what actually happened.

# Methodology III: (conditional) Average Treatment Effects

- Firm-level estimated treatment effects:

$$\hat{\alpha}_i = \hat{Y}_i^{SAM} - \hat{Y}_i^{SUM}.$$

- We compute average treatment effects (ATEs) by month and by subsamples defined according to firm characteristics, and we calculate bootstrapped standard errors.

- Our estimator will be **unbiased if the expected values of the prediction error of the SUM and of the SAM are the same in the relevant subsample (including the case in which they are both zero!)**.

- This estimator (Cerqua and Letta, 2020; Fabra et al., 2022):

$$\hat{\alpha}_i = Y_i^{observed} - \hat{Y}_i^{SUM}.$$

  is **unbiased only if expected value of the prediction error of the SUM is zero**

# Methodology IV: Effect Heterogeneity

- To uncover the possible heterogeneity of the effects, we use the **Sorted Effects method (Chernozhukov et al., 2018, 2020)**.
- First, we order the estimated individual specific treatment effects, we compute their **percentiles** and represent them graphically.
- Second, we **classify firms** as highly affected and weakly affected by COVID-19 according to whether their estimated individual effects are lower(greater) than the 25th (75th) percentile of the distribution of the estimated treatment effects.
- Third, we test which are the **characteristics on which these two group of firms differ** on average (difference in means of firm characteristics).
- We use the **bootstrap to calculate standard errors** of the difference in means and we calculate **joint p-values that account for simultaneous inference** (cause we are simultaneously testing many hypotheses).

# ML Models

- **The prediction performance out of sample of our empirical models is of fundamental importance because our identification strategy is based on the ability to reconstruct a counterfactual that is in practice out of sample, because it is unobserved**.
- Our approach recognizes that this is a complex task because
  - we have a very high number of potential explanatory variables
  - the existence of complex interdependencies between firms, and products and destinations that are difficult to know ex-ante.
- In such a situation, an approach that is based on the maximization of the accuracy of in-sample predictions will be prone to overfitting.
- Instead, ML techniques have been shown to constitute the best way to perform out of sample predictive tasks.
- We compare four different models: *Logit, Logit-Ridge, Logit-LASSO, and Random Forest (RF)*.

# Model Selection

- We focus on various statistics summarizing the predictive power of the models.
  - **Root Mean Squared Error:** The closer to 0, the better
  - **Area Under the receiver operating Curve (AUC):** Varies between 0.5 and 1, where 0.5 means that we predict randomly and 1 that the model predicts correctly all the individuals. AUC

- **Table 1** reports the accuracy of the estimates obtained studying the **probability of exporting in 2019 for the population of 2018 exporters by using cross-validation** (i.e., we use 5 folds and we build the prediction for the observations in each fold by learning in the other four folds).

- In this setting, **Logit-LASSO and RF models are the best performers**.

# SUM Models Performance in 2018/19

Table 1: Goodness of Fit for SUM in 2018/19

| | AUC | | | | RMSE | | | |
|---|---|---|---|---|---|---|---|---|
| | Logit-LASSO | Logit-Ridge | Random Forest | Logit | Logit-LASSO | Logit-Ridge | Random Forest | Logit |
| 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 |

# SUM Models Performance in 2019/20

- The models of Table **2** are also estimated using firms observed in 2018, their characteristics in 2018 as explanatory vars and their observed outcome in 2019.
- However, **these models are tested using the set of exporters of 2019 and their observed outcome in 2020**.
- **If the functions representing the relationship between explanatory variables and the outcome in absence of the pandemic are sufficiently similar for the pre-pandemic year and 2020, we expect that the accuracy in the first three months of 2019 and 2020 to be similar**.
- Indeed, during these months, the accuracy of Logit-LASSO and RF remains unchanged, as expected, compared to the accuracy obtained in Table 1.
- As expected, after April, the accuracy obtained in Table 2 is lower because it refers to the ability of a model trained without using any COVID-19 information to predict outcomes under a COVID-19 shock scenario.

Table 2: Goodness of Fit for SUM in 2019/20

| | AUC | | | | RMSE | | | |
|---|---|---|---|---|---|---|---|---|
| | Logit-LASSO | Logit-Ridge | Random Forest | Logit | Logit-LASSO | Logit-Ridge | Random Forest | Logit |
| 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 |

# SAM Models Performance in 2019/20

- Models in **Table 3** are **trained and tested with the universe of exporters in 2019 and their observed outcomes in 2020**.
- 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.

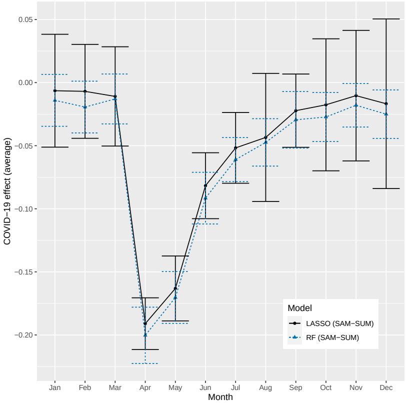
# SAM Models Performance in 2019/20

Table 3: Goodness of Fit for SAM in 2019/20

| | AUC | | | | RMSE | | | |
|---|---|---|---|---|---|---|---|---|
| | Logit-LASSO | Logit-Ridge | Random Forest | Logit | Logit-LASSO | Logit-Ridge | Random Forest | Logit |
| 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 |

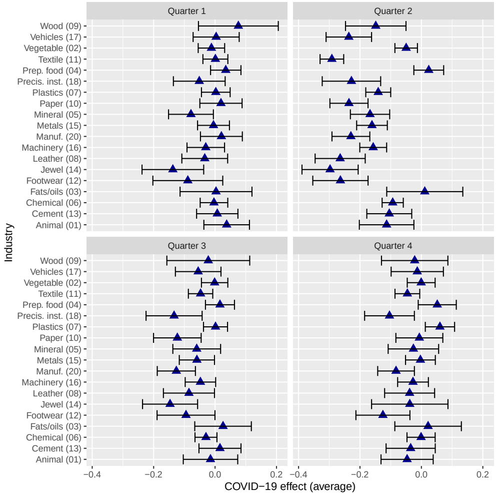
# Average Treatment Effect by Month

We use the Logit-LASSO predicted probabilities to estimate the average monthly effect of the COVID-19 shock as the monthly average of $\hat{\alpha}_i = \hat{Y}_i^{SAM} - \hat{Y}_i^{SUM}$.




- **In-Time Placebo** Suppose during first months of 2020 firms are not affected by COVID-19 shock (**lockdown March 25**) → Comparing SAM and SUM predictions is a falsification test (Abadie et al., 2015).

# Average Treatment Effect by Industry

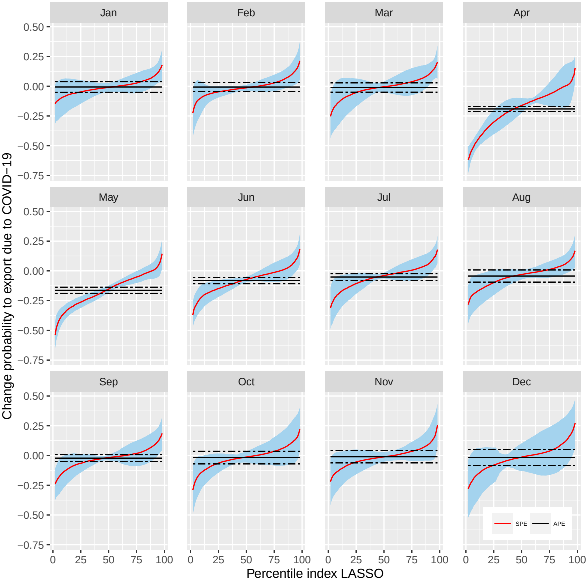



- The effects at the industry level are negative in general, but there are industries more affected than others. (Q1 effects by ind. are placebos for CATEs)

# Sorted effects by month

- The next figure shows the estimated **Sorted Partial Effects (SPE) by month**, which are just the **percentiles of the estimated individual treatment effects**, and 95% confidence intervals with blue bands (in black as a reference the average partial effects, APE=ATE).
- The main result is that we find **significant treatment effect heterogeneity just for the months of April, May** and, to a lesser extent, June, when statistically significant negative values are reported just in the left tail of the distribution.
- Instead, starting from July the confidence intervals of the SPEs intersect those of APE.
- Very importantly, we can also observe how **the SPEs coincide with the APEs in pre-pandemic months**, suggesting that our methodology is robust **also in the tails of the distribution of treatment effects**.

# Sorted effects by month



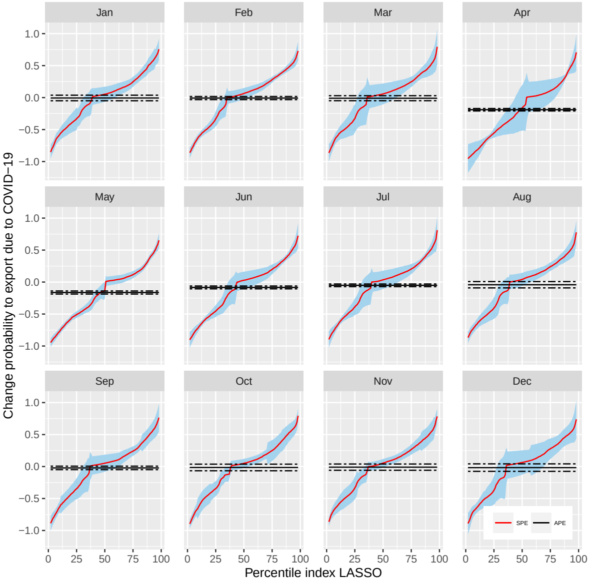
# Sorted effects by month: alternative estimator

- This is not true when using the alternative estimator (Cerqua and Letta, 2020; Fabra et al., 2022) **using the observed outcomes in 2020**:

$$\hat{\alpha}_i = Y_i^{observed} - \hat{Y}_i^{SUM}.$$

- Though the monthly ATE is almost the same, this method finds statistically **significant effects in both tails of the distribution when COVID-19 is absent**.
- Therefore, **firms in tails would be misclassified** as very positively or very negatively affected.
- This is due to the fact that **in tails the prediction error of $\hat{Y}_i^{SUM}$ is not zero on average**, but it turns out to be similar to the one of $\hat{Y}_i^{SAM}$!
- Therefore by taking differencs between $\hat{Y}_i^{SAM}$ and $\hat{Y}_i^{SUM}$ we wash out the estimation error.

# Sorted effects by month: alternative estimator



# Determinants of TE heterogeneity

- Finally, in order to identify the determinants of treatment effect heterogeneity, in the following Table we focus on **the difference in means of the the main explanatory variables across the most and least affected groups** (according to whether their estimated individual treatment effects are lower than the first quartile or greater than the third quartile, respectively).
- 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 indicating whether a firm belongs to the group of the most affected firms (in left tail of the distribution of the effects, with negative effects).
- Then, we also provide the **difference in adjusted means** once we have controlled for firm sector and both for firm sector and month of the year.

# Determinants of TE heterogeneity

- The dependent variables that we consider to explore the sources of COVID-19 treatment effect heterogeneity are firm characteristics observed in 2019 (the year before receiving the treatment): the industry, the means of transportation, the months when firms operate, the number of export destinations ($ND$), of import origins ($NO$), and of products ($NP$) exported.
- To explore to what extent treatment effect heterogeneity depends on the exposure of firms to COVID-19 through their activities on on international markets, we also consider as dependent variables the weighted Containment Stringency Indexes that exporters face when exporting and importing.

# Determinants of TE heterogeneity

| Outcome variable | $\beta^1_{1,f}$ | $\beta^2_{1,f}$ | $\beta^3_{1,f}$ |
|---|---|---|---|
| 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 | $-0.0621^{**}$ | |
| ND | -0.1990 | -0.1640 | -0.2480 |
| NO | -1.7470 | $-1.9820^{***}$ | $-2.4440^{**}$ |
| 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 | | | ✓ |

# Determinants of TE heterogeneity

- **The most affected exporters (those located in the left tail** of the estimated individual-level treatment effects distribution) experienced **a decrease in the probabilities to export between** 27.9 **p.p.and** 31.3 **p.p. lower** than the one experienced by the least affected firms (those located in the right tail).
- **The share of Textile firms among the most affected 2019 exporters is** 16 **p.p. higher** with respect to the one estimated for the group of the least affected firms. Similarly, we find the presence of 2.9 p.p. more wood exporters among the most affected than among the least affected firms.
- There are **more exporters using the air among the most affected** than among the least affected firms. However, there are less exporters using the sea to ship goods among the most affected than among the least affected firms.

# Determinants of TE heterogeneity

- **We do not find evidence that ex-ante diversification on the export side helps** to face a shock of this kind, as we can evince from the estimated parameters associated to *ND*, *NP*, and, in the first column, to *NO*.
- However, once we control for sector and therefore, inter alia, for the fact that some sector has relatively more diversification potential, **we find that the most affected Colombian exporters tend to import from** 1.98 **less countries in 2019 than the least affected firms**.
- The **most affected Colombian exporters face on average a higher Export (Import) Containment Stringency Index** with respect to the one faced by least affected firms.
- Finally, **the least affected firms exported and, especially, imported more value in 2019** than the most affected firms. Therefore, Colombian exporters trading in larger volumes (in value) are more resilient under a COVID-19 scenario.

# Concluding remarks

- In this paper, we show that causal ML can be a powerful tool for estimating ATE and investigating TE 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.
- While this method is specifically designed for analyzing the heterogeneous impacts of economy-wide shocks, there exists **potential utility in employing this approach also in less extreme situations where policies or shocks may exhibit unobservable spillovers that are challenging to model in advance**.
- In such contexts, our empirical framework proves advantageous in detecting potential heterogeneous indirect effects, as it **circumvents the need for a priori identification of a control group of untreated units**.

# Concluding remarks

- 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.
- By analyzing the estimated treatment effect distribution, we reveal that these average results hide considerable heterogeneity. For example, 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 and compare their characteristics using the SPE methodology.
- We emphasize how the integration into global value chains on the import side, both in terms of the number of countries from which a firm sources and the value of imports, is an important factor of resilience for exporters facing the COVID-19 shock.

Thanks for your attention

# Appendix

# App. 1: Empirical Strategy

Back

- We denote the potential outcome and the regressors under the scenario $d \in \{0, 1\}$ for firm $i$ at time $t$ as $Y_{it}^{d}$ and $X_{it}^{d}$, where $d$ is an indicator variable for the presence of COVID-19.
- The first step of the analysis is to estimate the counterfactual outcome in 2020: $Y_{i,2020}^{0}$.
- In particular, we will use the outcomes and covariates observed in 2018 and 2019 to reconstruct $Y_{2020}^{0}$ under the following assumptions (we omit $i$):

# App. 1: Empirical Strategy

(i) Both covariates and outcomes of 2018 and 2019 are not affected by the pandemic:

$$Y_t = Y_t^0 = Y_t^1, \quad X_t = X_t^0 = X_t^1 \quad \text{for} \quad t = 2018, 2019. \tag{1}$$

(ii) 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 absence of the pandemic such that $\mathbf{E}[Y_t^0|X_{t-1}^0] = f_t^0(X_{t-1}^0)$.

Under (i), for $t = 2019$ we have that $Y_{2019} = f_{2019}^0(X_{2018}) + u_{2019}^0$ such that $\mathbf{E}[Y_{2019}|X_{2018}] = f_{2019}^0(X_{2018})$.

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:

$$f_{2019}^0 = f_{2020}^0 = f^0 \tag{2}$$

# App. 1: Empirical Strategy

- Under the above assumptions, we can write $Y^0_{2020} = f^0(X_{2019}) + u^0_{2020}$, such that $\mathbf{E}[Y^0_{2020}|X_{2019}] = f^0(X_{2019})$, and we can use data on 2018 and 2019 to estimate $Y^0_{2019} = f^0(X_{2018}) + u^0_{2019}$ and retrieve $\hat{f}^0$.
- By applying this invariant estimated function to the covariates of 2019 we can obtain the predictions for the counterfactual (without COVID-19) outcome in 2020:

$$\hat{Y}^0_{2020} = \hat{f}^0(X_{2019}) = Y^0_{2020} - \overbrace{\mathcal{E}^0_{2020}(X_{2019})}^{\textit{Prediction error}} - \overbrace{u^0_{2020}}^{\textit{Orthogonal error}} \tag{3}$$

# App. 1: Empirical Strategy

- In general, the estimated counterfactual outcome in 2020, $\hat{Y}^0_{2020}$, will not be a perfect estimate for $Y^0_{2020}$ because $\hat{f}^0$ will not be a perfect estimate of $f^0$ thus producing a prediction error, which in the formula above we have denoted with $\mathcal{E}^0_{2020}(X_{2019}) = f^0(X_{2019}) - \hat{f}^0(X_{2019})$, and because of the existence of other determinants of the outcome that are orthogonal to the covariates, which in the formula above are contained in $u^0_{2020}$.
- The inaccuracy coming from the estimation of $f^0$, that can vary according to a firm's characteristics $X_{2019}$, will be reduced by experimenting with different ML techniques and using the one associated with the best out-of-sample performance.

# App. 1: Empirical Strategy

- Finally, we obtain the $\hat{Y}^0_{2020}$ by estimating $Y_{2019} = f^0(X_{2018}) + u^0_{2019}$ on entire set of 2018 exporters (also in this case month by month) and, as shown in (3), applying the estimated function $\hat{f}^0$ to the set of 2019 exporters.
- 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) 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 those of $\hat{Y}^0_{2020}$ for 2020.

# App. 1: Empirical Strategy

- Following Cerqua and Letta (2020) and Fabra et al. (2020), we define as an estimator of the individual-specific COVID-19 effect $\alpha$ the simple comparison of the observed outcome under COVID-19 in 2020 with the estimated counterfactual outcome for a given firm:

$$\hat{\hat{\alpha}} = Y_{2020} - \hat{Y}^{0}_{2020}. \tag{4}$$

- Eq. (4) provides the full distribution of treatment effects.
- All the parameters of interest of the paper are obtained by computing (conditional) averages and quantiles of such distribution.

# App. 1: Empirical Strategy

- Starting from Eq. (4), by taking the expected value of the individual treatment effect $\hat{\hat{\alpha}}$ for those units with $X_{2019} = x_{2019}$, we can define the following estimator of the conditional average treatment effect (CATE; the average effect for those units with $X_{2019} = x_{2019}$)

$$\begin{aligned}\mathbf{E}[\hat{\hat{\alpha}}|X_{2019} = x_{2019}] &= \mathbf{E}[(Y_{2020} - Y^0_{2020}) - \mathcal{E}^0_{2020} - u^0_{2020}|X_{2019} = x_{2019}] = \\ &= \underbrace{\Delta(X_{2019} = x_{2019})}_{CATE} - \mathbf{E}[\mathcal{E}^0_{2020}|X_{2019} = x_{2019}] - \underbrace{\mathbf{E}[u^0_{2020}|X_{2019} = x_{2019}]}_{=0 \text{ by assumption}},\end{aligned} \tag{5}$$

where,

$$\Delta(X_{2019} = x_{2019}) = \mathbf{E}[Y_{2020} - Y^0_{2020}|X_{2019} = x_{2019}].$$

- Therefore $\mathbf{E}[\hat{\hat{\alpha}}_i]$ will identify the unconditional average treatment effect, $\mathbf{E}[\Delta(X_{2019})] = \Delta$, if on average the prediction error is zero: $\mathbf{E}[\mathcal{E}^0_{2020}] = 0$.
- The conditional average treatment effect, $\Delta(X_{2019} = x_{2019})$, will be identified by $\mathbf{E}[\hat{\hat{\alpha}}_i|X_{2019} = x_{2019}]$ if on average the prediction error will be zero in the relevant sub-sample: $\mathbf{E}[\mathcal{E}^0_{2020}|X_{2019} = x_{2019}] = 0$.

# App. 1: Empirical Strategy

- Now let's decompose the outcome observed in 2020 in presence of the pandemic, $Y^1_{2020}$, in a generic model or function $f^1(X^1_{2019})$, which represents the relationship between explanatory variables and the outcome during the pandemic, and other determinants of the outcome, $u^1_{2020}$, that are orthogonal to the covariates

$$Y^1_{2020} = f^1(X^1_{2019}) + u^1_{2020}, \quad s.t. \ \mathbf{E}[Y^1_{2020}|X^1_{2019}] = f^1(X^1_{2019}). \quad (6)$$

- Given that $Y^1_{2020} = Y_{2020}$ and $X^1_{2019} = X_{2019}$, then

$$Y_{2020} = f^1(X_{2019}) + u^1_{2020}, \quad s.t. \ \mathbf{E}[Y_{2020}|X_{2019}] = f^1(X_{2019}). \quad (7)$$

# App. 1: Empirical Strategy

- At this point, we can define an alternative estimator of the individual-specific COVID-19 effect $\alpha$ as the comparison of the predicted outcome under COVID-19 in 2020 with the estimated counterfactual outcome for a given firm:

$$\hat{\alpha} = \hat{Y}_{2020} - \hat{Y}^{0}_{2020}, \tag{8}$$

where $\hat{Y}_{2020} = \hat{f}^{1}(X_{2019}) = Y_{2020} - \mathcal{E}^{1}_{2020} - u^{1}_{2020}$. We call "Shock Aware Machine" (SAM) the model that we use to predict $Y_{2020}$ (and the predictions $\hat{Y}_{2020}$ themselves)

# App. 1: Empirical Strategy

- Starting from Eq. (8), by taking the expected value of the individual treatment effect $\hat{\alpha}$ for those units with $X_{2019} = x_{2019}$, we can define the following alternative estimator of the conditional average treatment effect (for those units with $X_{2019} = x_{2019}$)

$$\begin{aligned} \mathbf{E}[\hat{\alpha}_i | X_{2019} = x_{2019}] =& \mathbf{E}[(Y_{2020} - Y^0_{2020}) - (\mathcal{E}^1_{2020} - \mathcal{E}^0_{2020}) - (u^1_{2020} - u^0_{2020}) | X_{2019} = x_{2019}] \\ =& \underbrace{\Delta(X_{2019} = x_{2019})}_{CATE} - \mathbf{E}[\underbrace{(\mathcal{E}^1_{2020} - \mathcal{E}^0_{2020})}_{\Delta\mathcal{E}} | X_{2019} = x_{2019}] - \\ & \mathbf{E}[u^1_{2020} - u^0_{2020} | X_{2019} = x_{2019}]. \end{aligned} \tag{9}$$

- Therefore, $\mathbf{E}[\hat{\alpha}_i]$ will identify the unconditional average treatment effect, $\mathbf{E}[\Delta(X_{2019})] = \Delta$, if on average the difference in prediction errors is zero: $\mathbf{E}[\Delta\mathcal{E}] = 0$.
- The conditional average treatment effect, $\Delta(X_{2019} = x_{2019})$, will be identified by $\mathbf{E}[\hat{\alpha}_i | X_{2019} = x_{2019}]$ if on average the difference in prediction errors is zero in the relevant sub-sample: $\mathbf{E}[\Delta\mathcal{E} | X_{2019} = x_{2019}] = 0$.

# App. 1: Empirical Strategy

- Given the definitions of SUM and SAM, to simplify the reasoning in the following we will refer to Equations (4) and (8) respectively as

$$\hat{\hat{\alpha}} = Y - \hat{Y}_{SUM} = Y - SUM. \tag{10}$$

$$\hat{\alpha} = \hat{Y}_{SAM} - \hat{Y}_{SUM} = SAM - SUM. \tag{11}$$

- The assumptions behind these identification results are not directly testable as they are expressed in terms of the expected values of the prediction error $\mathcal{E}^0_{2020}$ that is a function of the unobservable counterfactual $Y^0_{2020}$.
- The next table distinguishes the five different scenarios concerning the values of $\mathcal{E}^0_{2020}$ and $\mathcal{E}^1_{2020}$ that are relevant in determining whether applying the statistic **T** to *YSUM* and *SAMSUM* is able to recover the corresponding treatment effect estimand (e.g., whether averaging the estimated individual treatment effects would recover the average treatment effect).

# App. 1: Empirical Strategy

| | $\mathbf{T}(SAM - SUM)$ | $\mathbf{T}(Y - SUM)$ |
|---|---|---|
| $\mathbf{T}[\mathcal{E}^1_{2020}] \neq 0$ and $\mathbf{T}[\mathcal{E}^0_{2020}] = 0$ | X | ✓ |
| $\mathbf{T}[\mathcal{E}^1_{2020}] = \mathbf{T}[\mathcal{E}^0_{2020}] = 0$ | ✓ | ✓ |
| $\mathbf{T}[\mathcal{E}^1_{2020}] = 0$ and $\mathbf{T}[\mathcal{E}^0_{2020}] \neq 0$ | X | X |
| $\mathbf{T}[\mathcal{E}^1_{2020}] = \mathbf{T}[\mathcal{E}^0_{2020}] \neq 0$ | ✓ | X |
| $\mathbf{T}[\mathcal{E}^1_{2020}] \neq \mathbf{T}[\mathcal{E}^0_{2020}] \neq 0$ | X | X |

Table 4: Identification of generic functions of the individual treatment effects, **T**, according to the corresponding value taken by the prediction errors.

# Models: Logit, Ridge, LASSO

- **Logit** estimates the parameters maximizing the following log-likelihood function:

$$l(\beta) = \sum_{i=1}^{n} [y_i x_i \beta - log(1 + e^{x_i \beta})]$$

- **Logit-Ridge** adds a $L_2$ penalty to $l(\beta)$, that shrinks the parameters towards zero, without actually setting any of them to zero:

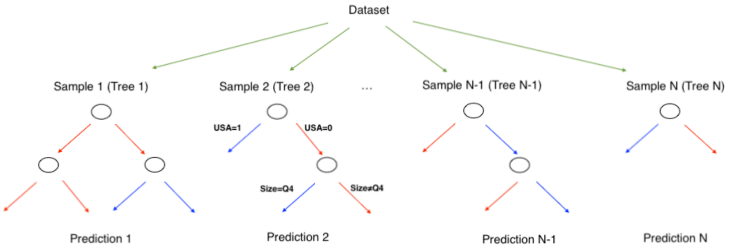
$$l(\beta) = \sum_{i=1}^{n} [y_i x_i \beta - log(1 + e^{x_i \beta})] - \lambda \sum_{j=1}^{p} \beta_j^2$$

- **Logit-LASSO** adds a $L_1$ penalty to $l(\beta)$, that forces some parameters to be exactly zero:

$$l(\beta) = \sum_{i=1}^{n} [y_i x_i \beta - log(1 + e^{x_i \beta})] - \lambda \sum_{j=1}^{p} \mid \beta_j \mid$$

# Models: Random Forest

- **RF** is composed by Random Trees. The final outcome of the RF is the average of the $N$ predictions.



# Models: Pros and Cons

- Logit as a benchmark model. Predicted performance is expected to be bad under large data sets or without a theoretical grounded model. Moreover, it is a high computational *cost* model. Possible problem of overfitting.

- Logit-Ridge is *faster* than Logit (for any fixed value of *lambda*). Good predictive performance when many variables of the model are relevant. But still possible overfitting problems when just few variables are relevant.

- Logit-LASSO has the benefit of *reducing the number of predictors* in the final model. Powerful when only a bunch of predictors have a *lot of predictor power*.

- RF is more robust to outliers. Moreover it takes into account all possible interactions, without specifying them. Every tree is independent of each other so RF *avoids overfitting*. However, RF has a high computational *cost*.

Back

| Pred | Actual |
|---|---|
| 0.83 | 1 |
| 0.72 | 0 |
| 0.44 | 1 |
| 0.31 | 0 |
| 0.61 | 1 |
| 0.55 | 0 |
| 0.91 | 1 |

| | Actual 1 | Actual 0 |
|---|---|---|
| Predict 1 | 3(TP) | 2(FP) |
| Predict 0 | 1(FN) | 1(TN) |

TPR = 3/(3+1) = 3/4    FPR = 2/(2+1) = 2/3