Estimation of Export Thresholds Distributions in French Sectors

Mattia Guerini[a,b,c], Patrick Musso[a], and Lionel Nesta[a,b,d]

[a]Université Côte d'Azur, GREDEG
[b]OFCE, SciencesPo
[c]Sant'Anna School of Advanced Studies
[d]SKEMA Business School

February 14, 2020

Abstract

To do.

Keywords: *To do.*
JEL Codes: *To do.*

✉ Mattia Guerini – mattia.guerini@univ-cotedazur.fr.

Mattia Guerini has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 799412 (ACEPOL). The usual disclaimers apply.

1 Introduction

In many economic decision problems, thresholds are unobservable barriers that might limit the action of economic actors. Thresholds are indeed well suited to represent in a very concise form the process trough which several economic decisions are taken: the presence of a minimum efficiency for a firm to participate to foreign markets; the presence of a minimum level of skills and capabilities for a firm to efficiently adopt a novel technology; the presence of a minimum amount of financial resources for a firm to repay the fix costs before the launch of a new investment project; the presence of a wage offer above the reservation wage for a worker to decide whether to search for a new job; etc. In all these cases, thresholds represent a minimum value below (above) which the economic actors decide to remain inactive (to become active) with respect to the economic activity of reference.

Modelling the participation problem by means of thresholds might be appealing from a theoretical viewpoint due to its simplicity. However, the major problem of such modelling choice is that thresholds are empirically unobservable to the social scientist. Even if an economic agent would truly take her participation decisions according to a threshold-overcoming problem, any external observer can only observe (i) the decision outcome and (ii) some individual characteristics of the decision maker. Thresholds are instead unobservable. In this paper, we show how to efficiently use the two available information in order to estimate the statistical properties of thresholds distributions. The advantage of our approach lays in the absence of strong requirements. Indeed, it only requires few working assumptions, which we explicitly present in the paper, together with the requirement that a monotonic relationship between the threshold support and the probability to participate exists (e.g. the larger the threshold the lower the probability to participate).

Our contribution is threefold. First, we propose a parametric Maximum Likelihood Estimation (MLE) approach to the discovery of the parameters characterizing threshold distributions. The problem can be derived by assuming a distribution of heterogeneous thresholds, conditional on the observation of the decision outcome as well as a critical variable of each individual. But for the MLE approach to be consistent and efficient, two underlying assumptions – one distributional and one behavioural – need to be correctly specified. We thus use stochastic MonteCarlo simulations in order to study the reliability of our approach when these assumptions are violated, and we broadly define the boundaries of its application. Second, we provide a primer empirical application to the problem of export thresholds. The application is naturally linked to the previous literature on international trade, where often it is assumed that firms decide to export whenever their efficiency level overcomes a homogeneous threshold (a.k.a. iceberg cost or productivity cut-off). This application on the one side allows us to extend the concept from homogeneous to heterogeneous cut-offs and on the other side allows us to rationalize the puzzle of overlapping productivities between exporters and non-exporters. In our application we also provide a horse-race study to search for the best explanation of empirical participation rates. We find that low-order moments of the threshold distributions are powerful tools to predict participation rates. Efficiency premiums alone are instead useless for the purpose. Third, we employ the estimates from our empirical exercise to investigate upon the possible of effects of policy shocks. In particular, we study the effects of an exchange rate shock and of a export subsidy shock. Overall, our results indicate that accounting for agents heterogeneity and for higher order moments, allows one to gain new relevant information.

Outline. The paper is organized as follows. Section 2 formally describes the economic problem under consideration and the tool employed to solve it. Section 3 presents the rationale and the results of extensive MonteCarlo exercises, which are performed to evaluate the statistical robustness of the proposed estimation strategy. An empirical application of our strategy to the export decision problem of French firms is presented in Section 4, together with an exercise aiming at the explanation of participation rates through the moments of the threshold distributions. Building upon these empirical results, Section 5 illustrates the practical importance of estimating thresholds for policy purposes. Section 6 concludes. The paper is complemented by Appendix A.2 which formally present the two distributions that we use throughout the paper, by Appendix A.1 which precisely describes the adopted MonteCarlo procedure and by Appendix A.3 which provides complementary details to the empirical application.

2 Econometric Strategy

The problem we tackle concerns a distribution of economic actors $i = 1, \ldots, N$ taking an economic decision whose outcome can be encoded as a binary variable $\chi_i \in \{0, 1\}$ representing the market participation of each individual. Each actor is characterized by an individual attribute θ_i that affects the decision outcome. This θ-attribute can be considered a single characteristic or a combination of several distinct features that ease or hinder the realization of a positive outcome $\chi_i = 1$. In particular, we assume that an actor takes the positive decision to participate only when the θ-attribute is sufficiently large so that it exceeds an individual threshold c_i, which can be interpreted as a barrier between the agent and the positive outcome.[1] This simple problem formally writes:

$$\begin{cases} \chi_i = 1 & \text{if} \quad \theta_i \geq c_i \\ \chi_i = 0 & \text{if} \quad \theta_i < c_i \end{cases} \tag{1}$$

An empirical social scientist is typically endowed with information about the decision outcome χ_i and the individual characteristics θ_i. But in most situations, the threshold variable c_i is a private information of the decision maker and is thus unobservable to the external observer.

The theoretical economic literature has most of the time assumed homogeneous across individuals thresholds – i.e. $c_i \sim \delta$, with δ representing the Dirac delta distribution (see Pissarides, 1974; McDonald and Siegel, 1986; Dixit, 1989, as early developers of such an approach).[2] Figure 1 describes this particular case. The grey area highlights all the agents taking a negative decision ($\chi_i = 0$) over the θ-attribute domain. This restriction implies therefore that a perfect separation of economic actors would arise. Only those agents which are able to overcome the unique threshold take a positive decision (e.g. efficient firms export while inefficient ones do

[1] Cases where thresholds are enhancing barriers rather then limiting ones are also possible in reality. The problem is simply the complementary of the one outlined in equation 1, but the approach that can be used to tackle this issue is equal.

[2] To our knowledge Cogan (1981) is the unique one that estimated heterogeneous thresholds in the labour market by means of a 2-steps strategy relying on a structural equations model.

not). But this implication is often contradicted by empirical evidence in several empirical domains – in particular about the efficiency of exporters (Bernard and Jensen, 2004) and about the efficiency of labour market bargaining processes (Alogoskoufis and Manning, 1991).

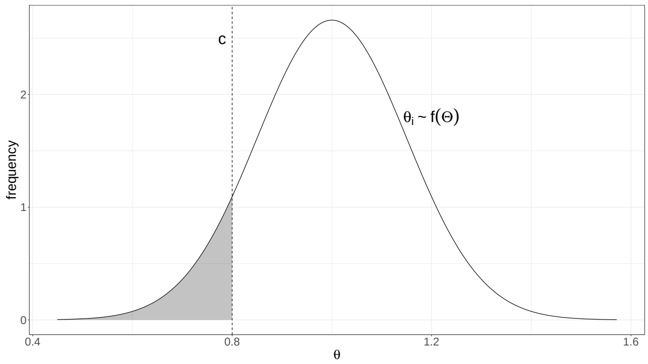

Figure 1: *Example of θ-attribute distribution with a common threshold c for all the individuals. Agents taking a negative decision are those whose θ-attribute lays in the grey highlighted interval.*

The empirical literature instead has often focused on the determination of the effects that some individual characteristic (the θ-attribute) have on the probability of taking the positive decision (e.g. Kau and Hill, 1972; Wei and Timmermans, 2008). This type of exercise can be done with simple Probit models at the cost of focusing only on the central moment of the threshold distribution. Also, this kind of approach is based on the assumption that one should always think in term of the threshold as a variable that is measurable over the θ-attribute domain. This approach allows one to match the fraction of actors with a positive (negative) decision outcome and to interpret the threshold as the probability of participation, conditional upon the θ-attribute.

By relaxing the restriction about homogeneity of thresholds, we instead observe the problem over the unobservable threshold domain, in relation to the measurable individual-specific θ-attribute. This perspective therefore inverts the x-axis and y-axis of Figure 1 and mirrors the individuals taking the negative outcome, as observable in Figure 2. An agent with a threshold c_i larger than her individual θ-attribute (thus displaying the negative outcome $\chi_i = 0$) would belong to the right side of the distribution. Her individual threshold would belong somewhere in the grey area of the right panel.

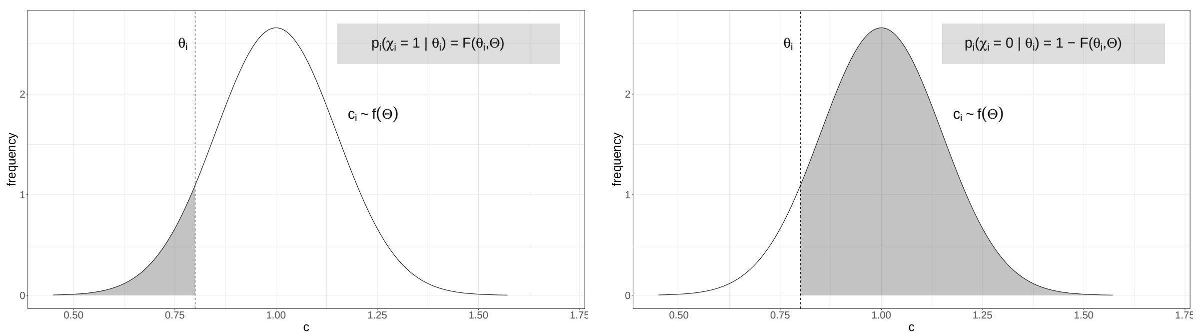

Figure 2: *Example of thresholds distribution. Left panel: an agent taking a positive decision has a threshold belonging to the grey interval. Right panel: an agent taking a negative decision has a threshold belonging to the grey interval.*

Our stance does not change the substantial mechanism behind the economic problem. The agents with an individual threshold lower (higher) than the individual θ-attribute will continue to take the positive (negative) decision. However, it demands for a different strategy to estimate the thresholds. Our interpretation requires a Maximum Likelihood Estimation approach, where the parameters to be estimates are the ones that define the whole thresholds distribution. In order to estimate these parameters, we simply need a prior description of the functional form taken by the unobservable threshold distribution.[3] In order to estimate the distribution of unobservable thresholds, we rely on the mild hypothesis that agents are heterogeneous in their θ-attribute as well as in their thresholds.

We can formally define the probability of an agent participating to the market as

$$p_i(\chi = 1|\theta_i) = F(\theta_i; \Omega) \tag{2}$$

and the probability of not participating as

$$p_i(\chi = 0|\theta_i) = 1 - F(\theta_i; \Omega), \tag{3}$$

the Likelihood function $L(\Omega)$ then takes the generic form:

$$L(\Omega) = \prod_{i=1}^{N} [F(\theta_i; \Omega)]^{\chi_i} \times [1 - F(\theta_i; \Omega)]^{1-\chi_i} \tag{4}$$

with F representing the cumulative density function of the probability distribution f. Two assumptions are required for our estimation strategy to work:

- **A1**: the vector of agent specific thresholds is a random variable following the density distribution $f(\Omega)$, where Ω is a vector of parameters characterizing the distribution f;
- **A2**: agents take the correct decision by comparing their individual threshold and θ-attribute, following the decision process defined by equation 1.

Under **A1** and **A2**, the Maximum Likelihood Estimation (MLE) is consistent and efficient. The vector of parameters Ω that characterize f can be easily recovered.

In particular, the first assumption concerns the functional form f of the thresholds. Since thresholds are unobservable, a-priori they can take any density function. The definition of f is therefore an important step that will impose constraints on the estimation of the threshold distribution. The second assumption concerns instead the information set available to the economic actors or to the external viewer and it is important in order to correctly locate the threshold of the economic agents with respect to their individual characteristic θ_i (e.g. the ones displaying a positive outcome shall have a threshold located to the left of it, as depicted in Figure 2).

However, since the *true* density function f of the thresholds is unobservable, and therefore unknown to the social scientist, an error might emerge when a functional form f different from the *true* one is assumed. Alternatively, an error might emerge because the economic actor might only have a limited information about either their own threshold c_i or their own θ-attribute; based on this limited information, they might take a bad decision (e.g. an agent might decide

[3] Notice that also in the traditional perspective, one has to assume a functional form for the threshold distribution (i.e. the Gaussian) and estimate the parameters that characterize it (i.e. the mean).

to participate while the rationale in equation 1 would have suggested not to). Finally, even if the agents would have a perfect information and would take the correct decision and even if the *true* functional form f is correctly assumed, it is the social scientist that might measure the θ-attribute of the individuals with some error; and therefore, estimate the threshold distribution based on a misspecified rule. All these three possibilities, might question the reliability of the estimation strategy. While the estimator is indeed consistent and efficient when **A1** and **A2** are valid, nothing is known about the consistency of the estimates when these two are violated. We thus conduct extensive MonteCarlo simulation exercises in order to investigate upon the robustness of the estimator under misspecified assumptions. These simulation exercises also allow us to select a functional form for f that is less discretionary: by minimizing the mean squared error (MSE) of the estimated moments over a sample of functional forms, we can find a locally optimal density function f^*.

An additional important feature of our method is that it allows one to compute the expected participation rate (EPR hereafter) starting only from (i) the observable empirical distribution of the θ-attribute (i.e. $\hat{\boldsymbol{\theta}}$) and (ii) the estimated cumulative distribution of the thresholds $F(\boldsymbol{\theta}, \Omega)$. This reads

$$\mathbb{E}[PR] = \frac{1}{N}\sum_{i=1}^{N} F(\theta_i, \Omega) \tag{5}$$

which is a simple average of the probability to participate, conditional on the cost function F and the productivity θ_i. Alternatively, if one has limited information about the empirical distribution of the θ-attribute and, for example, only has knowledge about the mean and the standard deviation (i.e. μ_θ and σ_θ), one can also use a fitted density function of the θ-attribute $g(\boldsymbol{\theta}|\mu_\theta, \sigma_\theta)$, which is conditional on the available moments information to estimate the EPR as follows:

$$\mathbb{E}[PR] = \int_{\underline{\theta}}^{\overline{\theta}} g(\boldsymbol{\theta}) F(\boldsymbol{\theta}, \Omega) \tag{6}$$

Notice however, that this second case requires an additional assumption about the functional form of the θ-attribute density function g.

3 MonteCarlo Simulations

In this section we describe results form MonteCarlo experiments aimed at studying the robustness of our approach. We begin in Section 3.1 with a perfect scenario, where both **A1** and **A2** are validated and we verify that our estimates are on average correct. In Section 3.2 we move toward imperfect scenarios, testing the robustness of the estimation when at least one of the two assumptions is false. Appendix A.1 describes in detail the simulation process for the MonteCarlo exercise.

3.1 Perfect scenario

Following the description put forward in Appendix A.1, we fix the number of firms to $N = 10000$ and we set the θ-attribute density $g \sim \mathcal{N}(1, 0.15)$.[4] We have then simulated the het-

[4]We have also verified that our results are qualitatively robust with respect to alternative specifications of g. We have tested $g \sim \mathcal{U}(min, max)$, $g \sim \mathcal{B}(\alpha, \beta)$, $g \sim \mathcal{P}(min, \alpha)$ where $\mathcal{U}, \mathcal{B}, \mathcal{P}$ represent respectively the Uniform, the

erogeneous threshold distribution $\boldsymbol{c}$ as extracted from a Gamma distribution with shape and scale parameters α and β – i.e. $f \sim \Gamma(0.8, 0.9)$ – and we have computed the vector of decision outcome $\boldsymbol{\chi}$.

Using only the limited information $\boldsymbol{\theta}$ and $\boldsymbol{\chi}$ we have then applied our MLE approach and we have estimated the vector of parameters $\Omega = \{\alpha, \beta\}$.[5] The distribution of the estimates for the shape and scale parameters over $M = 200$ MonteCarlo simulations are depicted in Figure 3.

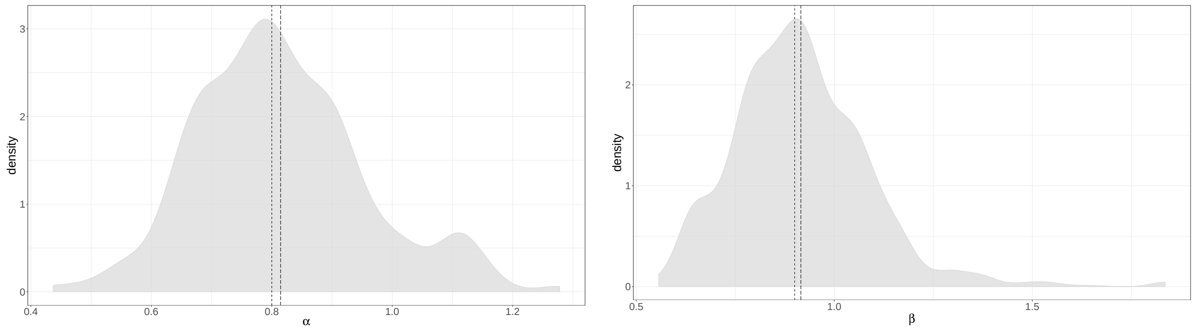

Figure 3: *Distribution of the $M = 200$ MonteCarlo estimates of shape and scale parameters. The true average is represented by a vertical line with short dashes. The estimated average is represented by the vertical line with long dashes.*

Our estimation strategy is able of correctly estimate on average the true parameters. As a matter of fact, a simple t-test does not reject the null hypothesis of equal mean. However, the MLE approach here proposed can somehow over- or under-estimate the true parameter by a quantitatively important magnitude. This is due to the presence of a strong relation between the estimated parameters α and β. As observable in Figure 4, there is indeed a clear non-linear and negative functional relationship linking the estimates of the two parameters; when one of the two parameters is under-estimated with respect to the true value, the other is over-estimated.

We believe that this is a positive feature for our estimation exercise in general. One shall indeed always remember that the objective of our research is not the correct estimation of the values of the parameters Ω. But the understanding of the shape of the functional form f. Therefore, the depicted non-linear relation between α and β simply signifies that there are a multiplicity of parameters combinations that allow us to recover a shape of the density function f sufficiently close to the true one.

As an example, in Figure 5 we report three examples of Γ distributions generated with three different parametrizations taken from points sourced in the scatter plot of Figure 4. There is a high similarity between the three distributions. Indeed the whole set of points following the non-linear pattern between on the α, β plane allows one to derive the same distribution. This means that, for our purpose, we shall not be too much worried about making errors in estimating one of the parameters characterizing f, as long as there exist a compensation effect with respect to the other parameter.

Beta and the Pareto type-II distributions. Results are available from the authors upon request.

[5] Notice that the first 4 moments of the Γ distribution are characterized as follows: $\mu_\Gamma = \alpha\beta$, $\sigma^2_\Gamma = \alpha\beta^2$, $sk_\Gamma = 2/\sqrt{\alpha}$ and $k_\Gamma = 2/\alpha$.

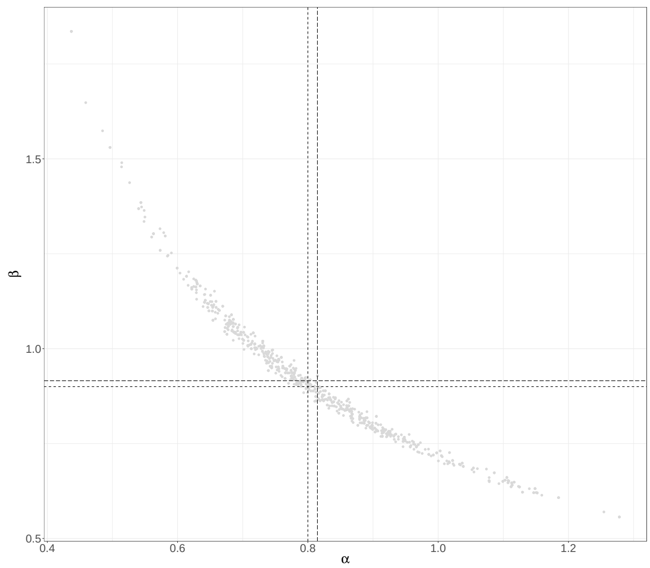

Figure 4: *Scatter plot of the estimated shape and scale parameters over $M = 200$ MonteCarlo simulations. A clear correlation structure is present. The true averages are represented by the lines with short dashes. The estimated averages are represented by the lines with long dashes.*

3.2 Imperfect scenarios

The second step consists in the evaluation of the estimation approach whenever each of the two underlying assumptions is violated.

3.2.1 Violation of A1

The first assumption concerns the selection of the threshold density function f, characterized by a vector of parameters Ω. The definition of f is indeed a necessary step whenever using a parametric MLE approach. But any probability distribution is suitable a-priori. Since thresholds are unobservable, any choice of f would only be an "educated guess". Given this a-priori uncertainty, understanding how much one can be mislead when the assumption **A1** is violated is of crucial importance. Additionally, our simulation exercise also guides us in the selection a probability density function f that is sufficiently flexible to take a variety of shapes and to adapt from case to case. The guidelines of the algorithm followed for this MonteCarlo exercise are presented in the Appendix A.1.

This exercise has been performed with $N = 10000$ firms, for $M = 200$ MonteCarlo runs and comparing two alternatives for the likelihood function $\mathcal{F} = \{\mathcal{N}(\mu, \sigma), \Gamma(\alpha, \beta)\}$, representing our assumption about the cost distribution. For the true distribution of the costs $\boldsymbol{c}$ we have also used either a Gaussian or a Gamma distribution. Additionally, for the Gamma case we further discriminate between two possibilities. One in which the distribution is almost symmetric and one in which the distribution is visually skewed.[6] Combining all the possible selection of f and $\boldsymbol{c}$ we obtain six different cases with assumption **A1** being violated in three occasions. To evaluate the exercise, we estimate the parameters, defining the moments of the distribution, and then compare the mean squared errors for the first four moments. This allows us to under-

[6]By construction the Gamma cannot be perfectly symmetric. However we have selected a parametrization such that the skewness is low.

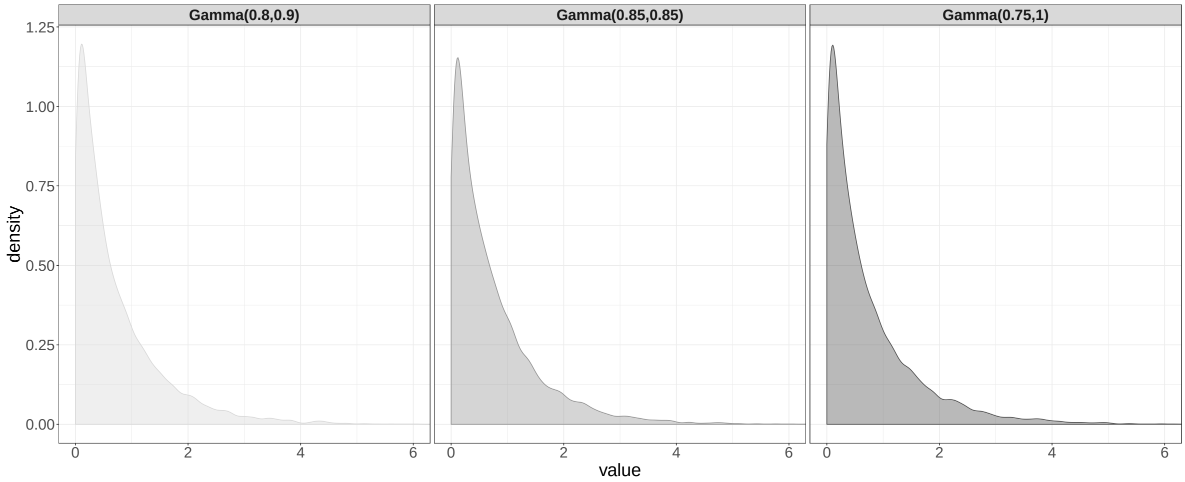

Figure 5: *Example of three distinct threshold distributions. For each distribution, the parameter vectors Ω has been sampled from points in Figure 4.*

stand which assumed cost distribution, between the Gaussian and the Gamma is safer. Since the threshold distribution is empirically unobservable, understanding which distributional assumption allows one to make fewer mistakes, even when starting from the wrong assumption, is therefore relevant for any empirical application.

Figure 6 presents the estimates of the average over the 200 MonteCarlo simulations in the different scenarios. Two are the interesting information provided by this figure. First, coherently with the previous subsection we again notice that when the assumption **A1** is satisfied, the estimation strategy works properly. The average of the true threshold distribution $\boldsymbol{c}$ is correctly estimated. Second, when the assumption is violated, both the maximum likelihood estimates (the Gaussian in the 1^{st} row, central and right panel; the Gamma in the 2^{nd} row, left panel) converge to a biased estimate of the average threshold distribution. This is due to a trivial property. The Normal distribution is symmetric by construction and cannot properly recover a cost distribution that is skewed, such as the Gamma (even when this skewness is low). Oppositely, the Gamma distribution is asymmetric by construction and cannot recover a symmetric cost distribution, like the Gaussian one. An additional qualification is needed. While for the Gaussian MLE a simple t-test ($H_0 : \hat{\mu_c} = \mu_c$) repeated for each MonteCarlo simulation, always reject the null hypothesis when the cost is truly Gamma distributed. For the Gamma MLE case instead, the t-tests never reject the null hypothesis even when the costs are Normal. This is a further indication for the better performance of the Gamma assumption.

The complete summary of results are presented in Table 1. We compare the Mean Squared Errors (MSE) of the first four moments for each of the six scenarios. As expected, each distributional assumption of $\boldsymbol{c}$, works better than the alternatives when the assumption is satisfied. For example, when the costs are truly Gaussian, the Normal MLE outperforms the Gamma MLE. But for the most interesting case of a violated **A1** assumption, we record an much better performance for the Gamma MLE. Comparing the 1^st row with the 4^th one, we observe that the misspecified Gamma MLE performs worse than a correctly specified Gaussian by a factor of 3.5 (on the estimation for the average). Oppositely, pairwise comparing the 2^nd and the 5^th rows as well as the 3^rd and the 6^th ones, we record that the misspecified Gaussian performs

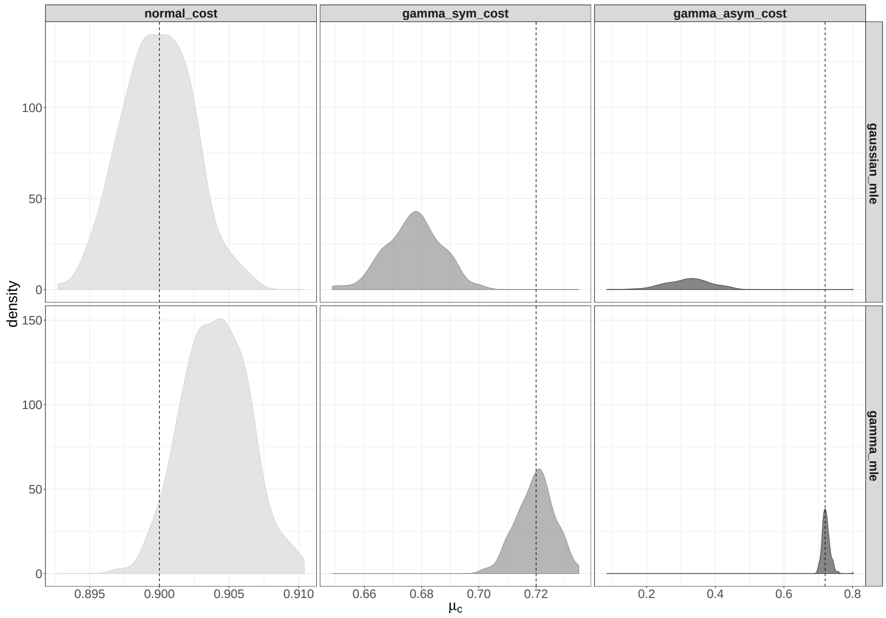

Figure 6: *Distribution of the MonteCarlo averages estimated over the different scenarios. The dashed vertical line represents the true average.*

worse than a correctly specified Gamma by factors of 45.1 and 1051.3 (on the estimation for the average). The relative magnitudes of the errors when a Gaussian distribution is incorrectly assumed are therefore much larger than the ones we register for a Gamma that is incorrectly assumed.

Assumed c	**True** c	**A1**	MSE_{μ}	MSE_{σ^2}	MSE_{sk}	MSE_{k}
$\mathcal{N}$	$\mathcal{N}$	✓	0.00000630	0.000000923	0.00	0.00
$\mathcal{N}$	Γ_s	✗	0.00193	0.000576	0.444	0.444
$\mathcal{N}$	Γ_a	✗	0.164	0.190	5.00	56.25
Γ	$\mathcal{N}$	✗	0.0000219	0.00000213	0.104	0.0246
Γ	Γ_s	✓	0.0000428	0.00000930	0.000506	0.00204
Γ	Γ_a	✓	0.000156	0.0183	0.0394	1.84

Table 1: *Mean Squared Errors of the estimated first four moments over the different scenarios.*

We therefore conclude that among the two alternatives tested, the Gamma distribution is the safest. This is due to the flexibility of such density function: variations in the estimates of the shape and scale parameters allow one to cover a wider set of shapes and functional forms.[7] This characteristics makes the Gamma more appropriate both when asymmetry and excess kurtosis are present and, more in general, when no prior knowledge about some of the statistical properties that characterize the distribution under investigation exist.

3.2.2 Violation of A2

The second assumption of our estimation strategy concerns the ability of the economic agents to follow the simple rule specified in equation 1. We acknowledge that, two possible types of errors can arise in such a framework. We define a *sorting error* a situation in which agent i,

[7] Details on the functional forms of the two distributions are available in Appendix A.2.

characterized by $\theta_i > c_i$ decides for the negative outcome χ_i. This can happen because agent i has bounded rationality and, for example, she has an imperfect evaluation of her threshold c_i. This leads to an error in the decision. With respect to Figure 2, a *sorting error* implies that by observing $\chi_i = 0$ we assign a threshold c_i to the right of the measured θ_i. But in reality the true threshold is at the left of it. We define instead a *measurement error* a situation in which we, as researchers, badly measure the observable θ-attribute of agent i. This error does not affect the decision of the agent, but might still affect the estimation due to a misspecification for the support of the threshold distribution. With respect to Figure 2, a *measurement error* implies some bias in the location of the θ-attributes for all the agents. Thus, we might use a support for our estimates that is misspecified. These two errors can also be present in an empirical exercise at the same time. Understanding how the estimation performance varies when **A2** is violated in any of these direction might therefore be a matter of practical applications.

To generate a *sorting error* in our MonteCarlo exercise, we modify the problem in equation 1 as follows

$$\begin{cases} \chi_i = 1 & \text{if} \quad \theta_i \geq c_i + \varepsilon_i^c \\ \chi_i = 0 & \text{if} \quad \theta_i < c_i + \varepsilon_i^c \end{cases} \tag{7}$$

where we assume that c_i represent the *true* threshold and $\varepsilon_i^c \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon)$ measures instead the bounded rationality of the agents, which is summarized by σ_ε. To simulate a *measurement error* instead, we keep the original decision problem in equation 1 unvaried, but we employ a noisy version of the θ-attribute when estimating the parameters characterizing the distribution of thresholds – i.e. the agents' characteristic observable by the researcher is $\theta_i + \varepsilon_i^\theta$, where $\varepsilon_i^\theta \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon)$ is a mean preserving spread. Also for these MonteCarlo exercises, the algorithmic guidelines are presented in Appendix A.1. While assumption **A2** is violated, for all these scenarios the assumption **A1** is valid instead. The underlying costs are distributed according to the Gamma law and the Gamma MLE is adopted.

In Figure 7 we present the results of the estimation of the Gamma parameters α and β for the three scenarios with *imperfect sorting* (IS) or/and *imperfect measurement* (IM). Each row represents a different scenario, as a function of the variance σ_ε, which increases from the left to the right in the columns. It is possible to observe that, as long as the error variance is relatively small, the estimates of the parameters characterizing the cost distribution $\boldsymbol{c}$ are correctly recovered independently of the type of the error type. In the *imperfect sorting* case (IS-PM) the estimation of parameters remain on average correct even when the degree of bounded rationality increases. In the *imperfect measurement* case instead (IS-PM), as the noise increases, the precision of the estimates worsens. When including both the imperfections (IS-IM case) the estimates do not seem to be worse than a case in which only the *measurement error* is present. However, for the last two cases, we can further notice that the correlation between estimates of α and β (cfr. Figure 4) is still present. A downward bias in the estimates for α is compensated by an upward bias in the estimates for β. Thus to properly evaluate the estimation one shall better look at the distributional properties. As a matter of fact, it is important to recall that the scope of our exercise is not the mere estimation of the parameters α and β. But the verification of the ability to recover information about the whole distribution $\boldsymbol{c}$.

To better quantify the performance in this respect, we present in Table 2 the MSE of the first

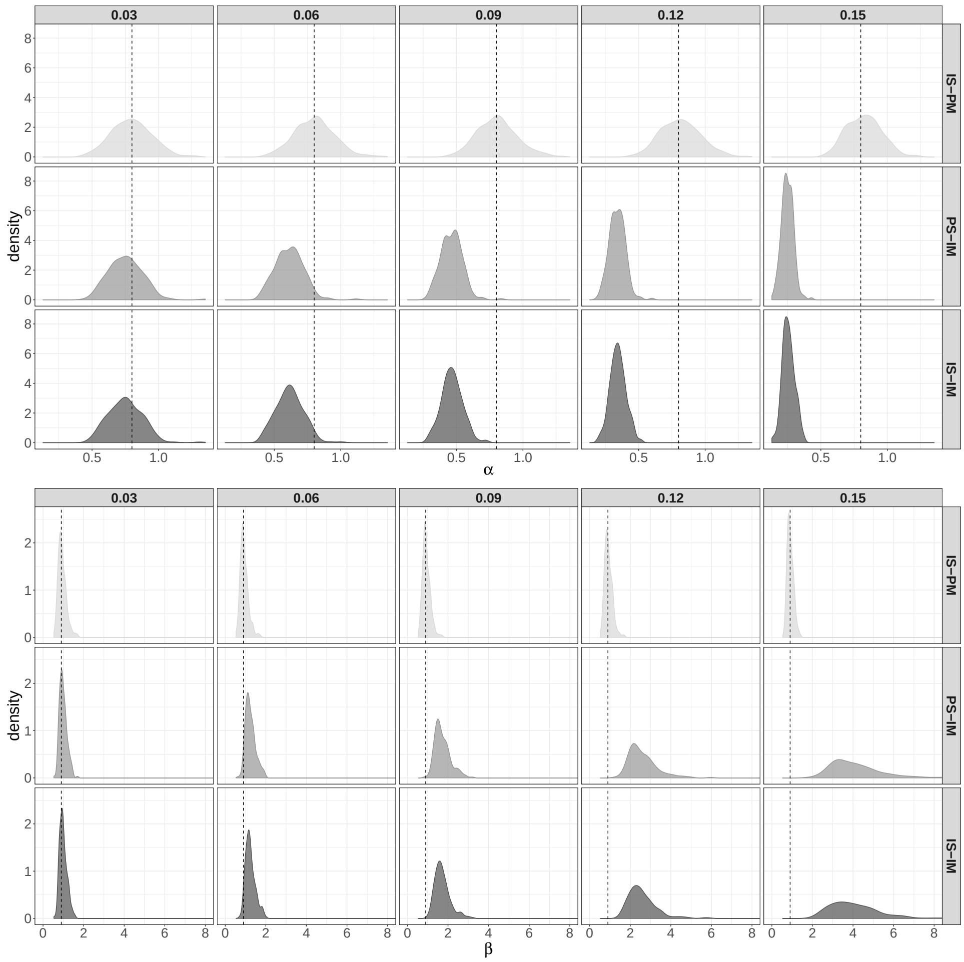

Figure 7: *Distribution of the MonteCarlo shape and scale parameters estimated over the different scenarios. The dashed vertical line represents the true average. PS: perfect sorting; IS: imperfect sorting; PM: perfect measurement; IM: imperfect measurement.*

four moments of the distribution for all the three scenarios where **A1** is violated.[8] The intuition provided by the parameter estimates in Figure 7 is confirmed. The presence of bounded rationality does not seem to affect the quality of the estimation, even when the variance increases. The presence of large measurement errors, oppositely, play an important role in providing biased estimates, especially about higher order moments (kurtosis in particular). Still, if the measurement error is relatively small ($\sigma_\varepsilon \leq 9\%$) the first two moments are estimated with a satisfactory precision. This, in general, allows us to conclude that in all the real world applications, a good measurement of the θ-attribute is important for precisely recovering the distribution of thresholds.

Collecting the results obtained with all the MonteCarlo simulations exercises, we conclude that our estimation strategy is robust when the assumptions **A1** and **A2** are satisfied. Additionally, the approach provides satisfactory results as long as their violations are relatively mild. Concerning the distributional assumption **A1**, we have concluded that unless some a-priori constraint can be imposed on the distribution of thresholds $\boldsymbol{c}$, it is safer to employ a flexible

[8]We also include in the table, for comparison, the case when **A1** is satisfied. For this case the estimates of the parameters α and β have instead already been reported in Section 3.1.

MSE	A2 (PS)	A2 (PM)	$\sigma_\varepsilon = 0\%$	$\sigma_\varepsilon = 3\%$	$\sigma_\varepsilon = 6\%$	$\sigma_\varepsilon = 9\%$	$\sigma_\varepsilon = 12\%$	$\sigma_\varepsilon = 15\%$
μ	✓	✓	0.0002					
σ^2	✓	✓	0.0183					
sk	✓	✓	0.0394					
k	✓	✓	1.8424					
μ	✗	✓		0.0002	0.0002	0.0002	0.0003	0.0003
σ^2	✗	✓		0.0262	0.0254	0.0237	0.0237	0.0141
sk	✗	✓		0.0495	0.0479	0.0452	0.0443	0.0311
k	✗	✓		2.4737	2.3762	2.1991	2.1235	1.4133
μ	✓	✗		0.0002	0.0006	0.0037	0.0201	0.0901
σ^2	✓	✗		0.0254	0.1130	0.6180	3.2625	19.4538
sk	✓	✗		0.0484	0.1675	0.6088	1.6319	3.4572
k	✓	✗		2.4911	9.6013	39.8571	126.9728	325.0632
μ	✗	✗		0.0002	0.0007	0.004	0.0228	0.1006
σ^2	✗	✗		0.0257	0.1128	0.625	3.3192	21.1340
sk	✗	✗		0.0482	0.1611	0.584	1.5518	3.2361
k	✗	✗		2.4797	9.1985	37.9831	119.3763	298.8161

Table 2: *Mean Squared Errors of the estimated first four moments over the different scenarios. PS: perfect sorting; PM: perfect measurement. Blank entries for inconsistent scenarios.*

distribution whose scales and shapes can easily vary, like the Gamma, which we have proposed as alternative to the Gaussian.[9] Focusing on the sorting and measurement assumption **A2** instead, we have concluded that a good measurement of the θ-attribute is of crucial importance for precisely estimating higher order moments of the distribution. In particular the importance of the correct measurement of the agents characteristics becomes more relevant when one is interested not only in the centrality measure (e.g. mean) but also in the dispersions, the asymmetry and the likelihood of extreme events. The sorting error instead, does not seem to create issues to our estimation strategy.

4 Empirical Application to International Trade

The export decision problem for a firm provides a perfect fit for our modelling framework. A firm exports whenever its efficiency (productivity) is larger than a firm specific threshold. Following the seminal contributions by Melitz (2003) and Melitz and Ottaviano (2008), the recent international trade literature have however modelled the export decision by means of a unique export threshold, identical for all the firms in a specific sector. Such an approach implies that in each sectors exporters efficiently self-select themselves. Only the most productive firms, whose productivity levels overcome the homogeneous threshold, are rationally deciding to enter the foreign market as exporters. But this implication is quite restrictive and clashes with the robust empirical evidence suggesting that there are both (i) firms characterized by high efficiency levels that decide not to export and (ii) firms characterized by low productivity levels that decide to export (Bernard and Jensen, 2004; Eaton et al., 2011; Impullitti et al., 2013).[10] A

[9]It is not the objective of this paper to find the “optimal” distributional assumption which is able to minimize a selected comparison criterion hence we did not focus on additional functional forms.

[10]The authors of the theoretical literature also recognize this limitation, but for analytical tractability motives, they cannot leave this assumption aside. Only recent versions of these models overcome this issue by accounting

possibility to reconcile this stylized fact with the empirical evidence is that of assuming that export thresholds are heterogeneous across firms.[11] The econometric approach outlined in section 2 allows us to estimate the density of this heterogeneous productivity cut-offs.

With this exercise, we contribute to the international trade literature by estimating the export cost distribution of a set of French manufacturing firms. We show indeed that productivity premiums alone are not able to predict the empirically observed participation rates. This implies that the productivity differential between exporters and non-exporters, is an insufficient statistics for the explanation of firms participation in foreign markets. Instead, by using the low order moments of threshold distributions estimated with our approach – across different combinations of sectors-years – we show that a Gamma threshold distributions is able to explain around 90% of the empirical participation rates. And this percentage of explained variance is larger than the one obtained by using a Gaussian density of threshold distributions. This also provides an indication for the presence of some degree of asymmetry in the empirical cut-offs distribution.

4.1 Data description

We use a panel database of French manufacturing firms covering the period 1990-2007 (EAE data) which covers the firms of at least 20 employees and with turnover higher than 5 millions Euro. The complete dataset has about 350 thousands observations. The relevant descriptive statistics for this dataset are presented in Table 3.[12] Around 73% of all the firms in our sample export. This is a relatively large participation rate, due to the fact that our dataset comprises only relatively large firms, which are more likely to export vis-à-vis small ones. Consistently with the previous economic literature, in all the industries – except for "Wood and paper" – the exporters are on average more productive than non-exporters as reported by the positive TFP premiums. In particular, exporters are on average about 4.2% more productive than non-exporter peers.

4.2 Estimating export participation thresholds

We apply our econometric strategy for estimating the distribution of export costs both at the aggregate manufacturing level as well as at specific sectoral levels.[13]

Results are reported in Table 4. We are a-priori agnostic about the best functional form of the density of the productivity cut-offs distribution, thus we perform our estimation exercise using both the Gaussian MLE (columns 2 to 4) and the Gamma MLE (columns 5 to 10). For the Gamma MLE, we also report the estimated median in order to provide a first insight about the possible presence of asymmetry in the distribution of thresholds. For most of the sectors

for product variety and heterogeneous product mix. Thresholds are equal within varieties but product mix are firm specific and generates firm specific productivity cut-off (Mayer et al., 2014).

[11] In such a framework one can interpret the thresholds as export costs measured over the productivity domain. In what follows therefore we might refer to thresholds also using the terms "export costs" or "productivity cut-offs".

[12] Notice that this statistics are obtained with the θ-attribute being measured as Total Factor Productivity. In the Appendix A.3 we report a similar table using the Apparent Labour Productivity (ALP) as θ-attribute.

[13] Results are similar and available upon request also at the year level for the whole manufacturing sector as well as for each combination of sector-year.

Industry Name	N. Firms	Participation Rate	TFP Premium
All manufacturing	357098	0.733	0.042
Automobile	9326	0.798	0.023
Chemicals	35473	0.836	0.024
Clothing and footwear	28055	0.672	0.143
Electric and Electronic components	14571	0.774	0.063
Electric and Electronic equipment	19323	0.754	0.065
House equipment and furnishings	23868	0.822	0.056
Machinery and mechanical equipment	62228	0.702	0.043
Metallurgy, Iron and Steel	61034	0.727	0.008
Mineral industries	15272	0.583	0.001
Pharmaceuticals	9169	0.913	0.034
Printing and publishing	29601	0.612	0.048
Textile	21885	0.798	0.080
Transportation machinery	5167	0.794	0.060
Wood and paper	22126	0.692	-0.015

Table 3: *Descriptive statistics of the participation rate by industry: number of firms in the sample, export participation rate and TFP premium of the exporters.*

indeed, the median is estimated to lay at the left of the average, indicating the presence of some degree of right-skewness and of a fatter than normal right-tail. The unique sector for which the skewness goes in the opposite direction instead is the Pharmaceutical industry, in which export participation rate is the highest. This result is rationalized by considering that the export costs for a large majority of the pharmaceutical products are relatively low. Cosmetics products and basic drugs are exported in large scale, and the larger costs of export might be fix, rather than variable. Only few pharmaceutical products might require particular and very costly arrangements.

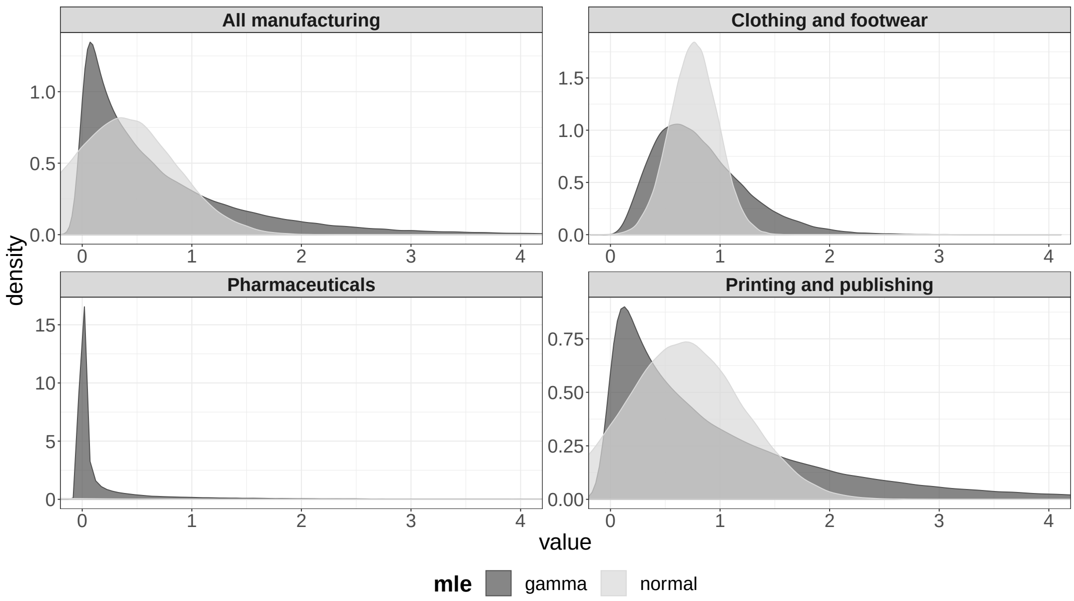

Figure 8: *Thresholds distributions for four selected sectors. Dark grey: thresholds estimated with a Gamma MLE distribution. Light grey: thresholds estimated with a Normal MLE distribution. Note: the Normal distribution has been truncated at 0.*

Given the parameter estimates of Table 4, we also graphically plot the export cost distri-

Industry Name	$\mathcal{N}_\mu$	$\mathcal{N}_\sigma$	$\mathcal{N}_\#$	Γ_α	Γ_β	Γ_μ	Γ_{med}	Γ_σ	$\Gamma_\#$
All manufacturing	0.364	0.487	5	0.783	0.939	0.735	0.460	0.345	8
Automobile	-0.141	0.671	5	0.319	1.983	0.632	0.310	0.627	10
Chemicals	-0.422	0.712	4	0.328	1.527	0.501	0.254	0.382	16
Clothing and footwear	0.763	0.218	6	3.782	0.217	0.822	0.749	0.089	14
Electric and Electronic components	0.449	0.346	5	1.251	0.522	0.653	0.488	0.171	11
Electric and Electronic equipment	0.465	0.365	4	1.225	0.559	0.686	0.511	0.192	15
House equipment and furnishings	0.265	0.378	4	1.010	0.548	0.554	0.386	0.152	8
Machinery and mechanical equipment	0.605	0.343	5	1.582	0.496	0.784	0.627	0.194	10
Metallurgy, Iron and Steel	-1.565	2.105	5	0.131	16.851	2.210	2	18.618	25
Mineral industries	-	-	-	-	-	-	-	-	-
Pharmaceuticals	-2.102	1.129	5	0.127	2.232	0.283	0.787	0.315	14
Printing and publishing	0.660	0.546	5	0.838	1.261	1.056	0.681	0.666	3
Textile	0.358	0.359	4	1.138	0.525	0.598	0.434	0.157	10
Transportation machinery	0.320	0.394	4	0.852	0.703	0.599	0.390	0.210	8
Wood and paper	-	-	-	-	-	-	-	-	-

Table 4: *Maximum likelihood estimation of participation thresholds distributions.* **Columns 2 and 3:** *estimated mean and standard deviation for the Gaussian MLE.* **Column 4** *number of iterations for convergence of the Gaussian MLE.* **Columns 5 to 9:** *estimated shape and scale parameters, mean, median and standard deviation for the Gamma MLE.* **Column 10:** *number of iterations for convergence of the Gamma MLE.* **Empty values:** *the MLE algorithm did not converge.*

butions under the two MLE assumptions by drawing a sample of observations randomly generated under that parametrization in Figure 8. Apart for the aggregate manufacturing industry, the three selected 2-digits industries are chosen because they are characterized by different shapes of the threshold distributions. Clothing and footwear is a case of quasi-symmetric costs, where the Normal and the Gamma MLE are, on average, quite similar. The pharmaceutical industry represents instead the case of a very skewed distribution. From the figure one can notice that while the Gamma is sufficiently flexible to capture peaks close to zero as well as the fatness of the right-tail, the Gaussian assumption seems restrictive. The normal has by construction no asymmetric behaviour and zero excess kurtosis. Hence, when using the Normal MLE, the distribution of thresholds results with a negative average and a very large standard deviation. Over the productivity domain, the normal distribution is almost flat.[14] The printing and publishing sector is instead representative of the aggregate manufacturing industry displaying an intermediate degree of skewness. In this case, both the Gaussian and the Gamma MLEs estimate positive average costs, however in order to be meaningful, the Normal distribution needs again to be truncated, not to enter over the negative domain.

A final remark from Table 4 concerns the two empty rows for “Mineral industries” and for “Wood and paper”. Both the Normal and the Gamma MLEs do not converge to a point in which the first derivative of the likelihood becomes null. This is due to a misspecification problem. Indeed, by looking at Table 3 one can notice that for the first industry the export premium is zero while for the second it is even negative. This, in terms of the decision problem outlined in equation 1, implies that our base assumption is not validated for these sectors. Indeed, in the “Wood and Paper” industry the decision problem is more likely to be reversed; for the “Mineral industries” the decision would be completely random instead, independent from the θ-attribute that we have taken here as the critical dimension affecting the firm export extensive margin decision.

[14]In the plots, we show the truncated Normal distribution with only positive values.

4.3 Explaining the export participation rates

At this stage we have an idea about the form of the distribution of thresholds $\boldsymbol{c}$ and of their sectoral specificities. But one might want to question the whole theory about the heterogeneity of export costs. Does it have some empirical relevance? Does it matter? Or, alternatively, is the productivity premium a sufficient statistics for the prediction of empirical participation rates? If, indeed, a perfect separation between high productivity exporters and low productivity non-exporters exists, then there is no need for the estimation of the $\boldsymbol{c}$ distributions. In that case, one can simply use a Probit model and estimate the optimal cut-off level $\tilde{\theta}$ of efficiency below which firms do not export. On top of that, even without a rejection of the heterogeneous thresholds theory, one can debate about the different estimations provided by the $\mathcal{N}$ and the $-$ MLE. In particular, now that we have estimated both of them, how can we discriminate between them? Some of our results, intuitively suggest that the $-$ alternative is more apt for coping with asymmetric and leptokurtic cost distributions. But are the third and fourth moments relevant?

In what follows we develop an exercise tackling these two issues. By estimating three different econometric models with participation rates as the dependent variables, we evaluate the explanatory powers of the variables therein included. More precisely, in the first stage we estimate the first three moments of the threshold distributions at the industry-year level, using the approach described above, to obtain the vectors of estimated moments $(\mathcal{N}_{\mu}, \mathcal{N}_{\sigma}, \Gamma_{\mu}, \Gamma_{\sigma}, \Gamma_{sk})$. For each industry-year pair, we also collect information about the export premium EP – measured as the difference between the average productivity of exporters and of non-exporters – as well as about participation rate PR – measured as the fraction of firms engaged in export activity.[15] In a second stage, we separately estimate the following models

$$PR_{i,t} = \begin{cases} \alpha + \beta_1 \; EP_{i,t} + \varepsilon_{i,t} \\ \alpha + \beta_2 \; \mathcal{N}_{\mu,i,t} + \beta_3 \; \mathcal{N}_{\sigma,i,t} + \varepsilon_{i,t} \\ \alpha + \beta_4 \; \Gamma_{\mu,i,t} + \beta_5 \; \Gamma_{\sigma,i,t} + +\beta_6 \; \Gamma_{sk,i,t} + \varepsilon_{i,t} \end{cases} \tag{8}$$

The first specification allows us to study whether export premia alone positively correlate with participation rates, answering the question about whether sectors where the differentials of productivity between exporters and non-exporters are higher, are also the sectors where the participation rate is larger, as implied by a homogeneous threshold assumption. The second and third specifications instead, allow us to evaluate the extent to which the key statistical properties of the export thresholds distributions explain the extensive margins. And they allow us also to compare the performances of the Normal and the Gamma MLE.

Results are reported in Table 5. Concerning the first model (column 1) we register that export premia alone are unrelated to participation rates. The estimated coefficient is not significant and the regression R^2 is extremely low. This suggest that one shall look to alternatives explanation to sectoral participation rates. The one proposed here is consistent with the heterogeneous thresholds theory. Estimates from different alternatives of the second and third

[15] This exercise is performed at a finer detail of the industrial classification (NAF114) in order to obtain more observation for each regression. Results are however robust with respect to all the three different industrial classifications available in the data (NAF36, NAF114 and NAF700). See Appendix A.3 for additional details.

	Dependent variable: Participation Rate (PR)							
	(1)	(2a)	(2b)	(3a)	(3b)	(3c)	(2b-fe)	(3c-fe)
β_1	0.212							
	(0.134)							
β_2		-0.065***	-0.166***				-0.034***	
		(0.007)	(0.007)				(0.008)	
β_3			-0.138***				-0.032***	
			(0.007)				(0.008)	
β_4				-0.395***	-0.457***	-0.582***		-0.514***
				(0.011)	(0.008)	(0.010)		(0.020)
β_5					0.033***	0.077***		0.073***
					(0.002)	(0.003)		(0.004)
β_6						-0.034***		-0.036***
						(0.002)		(0.003)
α	0.740***	0.774***	0.936***	1.033***	1.056***	1.186***	0.803***	1.144***
	(0.008)	(0.005)	(0.009)	(0.008)	(0.005)	(0.009)	(0.013)	(0.016)
FE_i							✓	✓
FE_t							✓	✓
Obs.	252	195	195	176	176	176	195	176
R^2	0.010	0.295	0.762	0.882	0.956	0.981	0.967	0.993

*** p<0.01, ** p<0.05, * p<0.1

Table 5: *Explanation of export participation rates. Number of observations varies due to the exclusion of sector-year combinations with less than 100 firms, to the lack of convergence of the MLE, to the trimming process of sectors with variance exceeding 10.*

models (columns 2· and 3·) describe indeed a significant and negative association between the average estimated cost and the participation rate. The higher the average cost, the lower the extensive margin. The standard deviation of costs has an ambiguous effect instead. When costs are estimated according to a Normal law, the variance also has a negative correlation with participation rates (column 2b). Looking instead at the costs estimated with a Gamma MLE, the dispersion measure registers a positive impact on participation rates (3b). This difference can be explained by the fact that the Gaussian case, being symmetric by construction tries to capture a fatter right tail by an increase in the second moment. For the Gamma case instead, allowing for asymmetric distributions, capture the right tail by means of a larger skewness, which in turn has a negative impact on participation rates (column 3c). The results are robust to the presence of sector and year fixed effects. Finally, the estimated R^2 allow us to discriminate between the Gaussian and the Gamma alternative. Without accounting for industry and time specific fixed effects indeed, the properties of the threshold distribution under a Normal setting explain up to around 75% of the total variance of participation rates. Under a Gamma case, the explained variance rises to around 95%. Our results therefore, suggest that the Gamma distribution, with asymmetry and fat tails, better captures the firm-level heterogeneity and is therefore more appropriate in describing the true underlying export threshold distributions.

4.4 Estimating and explaining entry and remaining rates

The previous two subsections have analysed the participation rate. But the ones and zero describing the outcome variable χ in our original problem (cfr. equation 1) for exporting and non-exporting firms respectively might have different natures and might suggest different decisions. In particular, one can further exploit the time dimension and condition the decision upon the export status of the previous period. Among the participating firms at time t, one can therefore more precisely distinguish between: (i) *entrants firms*, participating into the foreign market in t and not participating in period $t-1$; (ii) *remaining firms*, participating into the foreign market in t and also in period $t-1$. Therefore one can focus on the estimation of the threshold distributions for each specific sub-category.

Industry Name	$\mathcal{N}_\mu$	$\mathcal{N}_\sigma$	$\mathcal{N}_\#$	Γ_α	Γ_β	Γ_μ	Γ_{med}	Γ_σ	$\Gamma_\#$
All manufacturing	2.992	1.326	5	0.474	28.699	13.611	2	195.311	24
Automobile	-	-	-	-	-	-	-	-	-
Chemicals	3.789	2.237	5	0.295	124.271	36.697	2	2280.223	30
Clothing and footwear				1.400	2.048	2.868	2	2.936	7
Electric and Electronic components	1.692	0.541	30	1.318	1.966	2.592	1.972	2.548	11
Electric and Electronic equipment	-	-	-	-	-	-	-	-	-
House equipment and furnishings	1.699	0.640	5	0.977	3.001	2.932	2	4.399	4
Machinery and mechanical equipment	-	-	-	0.563	14.575	8.212	2	59.841	16
Metallurgy, Iron and Steel	5.010	2.729	3	-	-	-	-	-	-
Mineral industries	-	-	-	-	-	-	-	-	-
Pharmaceuticals			No sufficient number of entrant firms						
Printing and publishing	2.773	1.169	12	0.610	13.103	7.997	2	52.391	16
Textile	-	-	-	0.669	8.677	5.804	2	25.179	9
Transportation machinery			No sufficient number of entrant firms						
Wood and paper	-	-	-	-	-	-	-	-	-

Industry Name	$\mathcal{N}_\mu$	$\mathcal{N}_\sigma$	$\mathcal{N}_\#$	Γ_α	Γ_β	Γ_μ	Γ_{med}	Γ_σ	$\Gamma_\#$
All manufacturing	0.364	0.487	5	0.783	0.939	0.735	0.460	0.345	8
Automobile	-0.141	0.671	5	0.319	1.983	0.632	0.310	0.627	10
Chemicals	-0.422	0.712	4	0.328	1.527	0.501	0.254	0.382	16
Clothing and footwear	0.763	0.218	6	3.782	0.217	0.822	0.749	0.089	14
Electric and Electronic components	0.449	0.346	5	1.251	0.522	0.653	0.488	0.171	11
Electric and Electronic equipment	0.465	0.365	4	1.225	0.559	0.686	0.511	0.192	15
House equipment and furnishings	0.265	0.378	4	1.010	0.548	0.554	0.386	0.152	8
Machinery and mechanical equipment	0.605	0.343	5	1.582	0.496	0.784	0.627	0.194	10
Metallurgy, Iron and Steel	-1.565	2.105	5	0.131	16.851	2.210	2	18.618	25
Mineral industries	-	-	-	-	-	-	-	-	-
Pharmaceuticals	-2.102	1.129	5	0.127	2.232	0.283	0.787	0.315	14
Printing and publishing	0.660	0.546	5	0.838	1.261	1.056	0.681	0.666	3
Textile	0.358	0.359	4	1.138	0.525	0.598	0.434	0.157	10
Transportation machinery	0.320	0.394	4	0.852	0.703	0.599	0.390	0.210	8
Wood and paper	-	-	-	-	-	-	-	-	-

Table 6: *Maximum likelihood estimation of entry (top) and remaining (bottom) threshold distributions.* **Columns 2 and 3:** *estimated mean and standard deviation for the Gaussian MLE.* **Column 4** *number of iterations for convergence of the Gaussian MLE.* **Columns 5 to 9:** *estimated shape and scale parameters, mean, median and standard deviation for the Gamma MLE.* **Column 10:** *number of iterations for convergence of the Gamma MLE.* **Empty values:** *the MLE algorithm did not converge.*

A priori, one would expect that thresholds are higher for entrant firms than for remaining firms. Previous exporters indeed, have already paid fix and sunk costs related to the entry in a new market in the past and at period t, only have to pay for the variable part. The results that we obtain and present in Table 6 are consistent with this argument. We find that for remaining firms the average cut-offs are lower than the ones for the entrant firms. This result is valid across all the sectors where a comparison is feasible. And both when using Gaussian or Gamma MLE. Additionally, when using the Gamma MLE we notice that the thresholds for the entrant firms might be characterized by very large values of the scale parameter β. On one side this might be the occurrence of the identification problem already outlined in the simulation section above (see Figure 4); on the other side, it might reflect a specificity of the distribution of entrant firms. A graphical comparison between the threshold distributions for entry and remaining for the aggregate manufacturing sector is depicted in the additional material in Appendix A.3 (see Figure 11).

	Dependent variable: Entry Rate (ER)							
	(1)	(2a)	(2b)	(3a)	(3b)	(3c)	(2b-fe)	(3c-fe)
β_1	0.104							
	(0.162)							
β_2		-0.013***	-0.069***				-0.052***	
		(0.004)	(0.008)				(0.011)	
β_3			0.041***				0.027***	
			(0.006)				(0.007)	
β_4				-0.039**	-0.242***	-0.198***		-0.144***
				(0.016)	(0.027)	(0.017)		(0.042)
β_5					0.047***	0.017***		0.011
					(0.006)	(0.005)		(0.009)
β_6						0.189***		0.173***
						(0.020)		(0.037)
α	0.221***	0.331***	0.361***	0.323***	0.614***	0.349***	0.343***	0.281***
	(0.005)	(0.011)	(0.009)	(0.035)	(0.043)	(0.038)	(0.020)	(0.088)
FE_i							✓	✓
FE_t							✓	✓
Obs.	230	58	58	55	55	55	58	55
R^2	0.002	0.171	0.574	0.105	0.600	0.855	0.844	0.952

*** p<0.01, ** p<0.05, * p<0.1

	Dependent variable: Remaining Rate (RR)							
	(1)	(2a)	(2b)	(3a)	(3b)	(3c)	(2b-fe)	(3c-fe)
β_1	0.312***							
	(0.095)							
β_2		-0.049***	-0.164***				-0.044***	
		(0.006)	(0.005)				(0.004)	
β_3			-0.153***				-0.044***	
			(0.005)				(0.004)	
β_4				-0.430***	-0.476***	-0.637***		-0.379***
				(0.008)	(0.009)	(0.009)		(0.016)
β_5					0.023***	0.090***		0.050***
					(0.002)	(0.003)		(0.003)
β_6						-0.051***		-0.026***
						(0.002)		(0.002)
α	0.759***	0.794***	0.968***	1.074***	1.090***	1.262***	0.879***	1.078***
	(0.006)	(0.004)	(0.006)	(0.006)	(0.006)	(0.008)	(0.011)	(0.014)
FE_i							✓	✓
FE_t							✓	✓
Obs.	716	543	543	484	484	484	543	484
R^2	0.015	0.121	0.713	0.848	0.876	0.945	0.947	0.978

*** p<0.01, ** p<0.05, * p<0.1

Table 7: *Explanation of export entry (top) and remaining (bottom) rates. Number of observations varies due to the exclusion of sector-year combinations with less than 100 firms, to the lack of convergence of the MLE, to the trimming process of sectors with variance exceeding 10.*

Also at the entry and remaining levels, we can evaluate the above mentioned theories using the econometric models specified by equation 8. From Table 7 we observe that export premium alone are not capable of explaining any feature of the sector-year entry and remaining rates. For entrant firms the parameter is non significant. For remaining firms, even if the parameter is positive indicating the presence of some form of correlation between the two variables, the regression explanatory power is extremely low. Concerning the appropriateness of the $\mathcal{N}$ vis-à-vis the Γ distributions, we observe the same result already encountered for participation rates. For both entry and remaining rates – and independently on the presence of time and industry fixed effects – the Γ MLE seems to capture a larger portion of the variance. Again, this is particularly true when the regression accounts for the third moment, which is null by default in the Gaussian case. This confirms the presence of relevant asymmetric behaviour in the distribution of export thresholds.

5 Policy Implications

Up to here we have empirically estimated export thresholds distributions and verified their relevance for the explanation of participation, entry and remaining rates. We now bring the results stemming from our approach to a further level of analysis. We tackle a political economy question concerning the effect of an exogenous shock (e.g. an exchange rate or a policy shock) on the overall share of exporting firms. As a matter of fact the aggregate effect of a shock can differ according to the shape of the thresholds distributions, which describes the degree of heterogeneity between firms. We here focus on the effects of (i) a *cost-shock*, directly impacting on the thresholds $\boldsymbol{c}$; and (ii) a *theta-shock*, which affects instead the θ-attribute (i.e. the firm-level productivity).

To perform this exercise we assume each industry is populated by $N = 10000$ firms, characterized by log-productivities θ_i drawn from a normal distribution with mean and standard deviations coherent with our empirical estimates for the aggregate manufacturing sector and the other three selected industries, as reported in Table 8. We then assume that these sectors have a threshold distribution which can be either $\mathcal{N}$ or Γ, also characterized by the parameters estimated at the previous stage. We then draw an individual threshold for each firm. As a result we have three empirically grounded vectors representing the distributions for: (i) the θ-attribute; (ii) the thresholds coherent with a Normal distribution; (iii) the thresholds coherent with a Gamma distribution.

We finally assume that firms follow the simple behavioural rule specified in equation 1, and participate in the market only if their individual productivity is larger than their individual threshold. It is straightforward then to compute for both scenarios, the share of exporting firms. This represent the estimated participation rate. We begin by observing that the estimations provided by the Γ are much closer to the observed empirical values (cfr. Table 3) while the Gaussian estimates can be far off the empirical target. This provides another support for the importance of higher order moments not captured by the Normal distribution.

The policy exercise consists then in the application of either a *cost-shock* or a *theta-shock*. Formally the shocks affect the two individual vectors as follows:

$$x_i^s = x_i(1+\tau) \tag{9}$$

Industry Name	$\mathcal{N}_\mu$	$\mathcal{N}_\sigma$	Γ_μ	Γ_σ	θ_μ	θ_σ	$\mathbb{E}[PR_{\mathcal{N}}]$	$\mathbb{E}[PR_\Gamma]$
All manufacturing	0.36	0.49	0.74	0.83	0.98	0.16	0.88	0.73
Clothing and footwear	0.76	0.22	0.82	0.42	0.98	0.20	0.76	0.67
Pharmaceuticals	-2.10	1.13	0.28	0.79	0.98	0.19	1.00	0.91
Printing and publishing	0.66	0.55	1.06	1.15	0.98	0.18	0.71	0.61

Table 8: *Estimated parameters used as inputs for the policy exercises.*

where $x \in \{c, \theta\}$ represents the shocked variable and the superscript s indicates the same variable after having received the shock.[16] The *cost-shock* can be interpreted as a general subsidy provided by the government to abate the export costs paid in the domestic currency. The *theta-shock* instead can represent the effect of an increase in the real exchange rate that allows local firms to become more competitive vis-à-vis foreign ones.

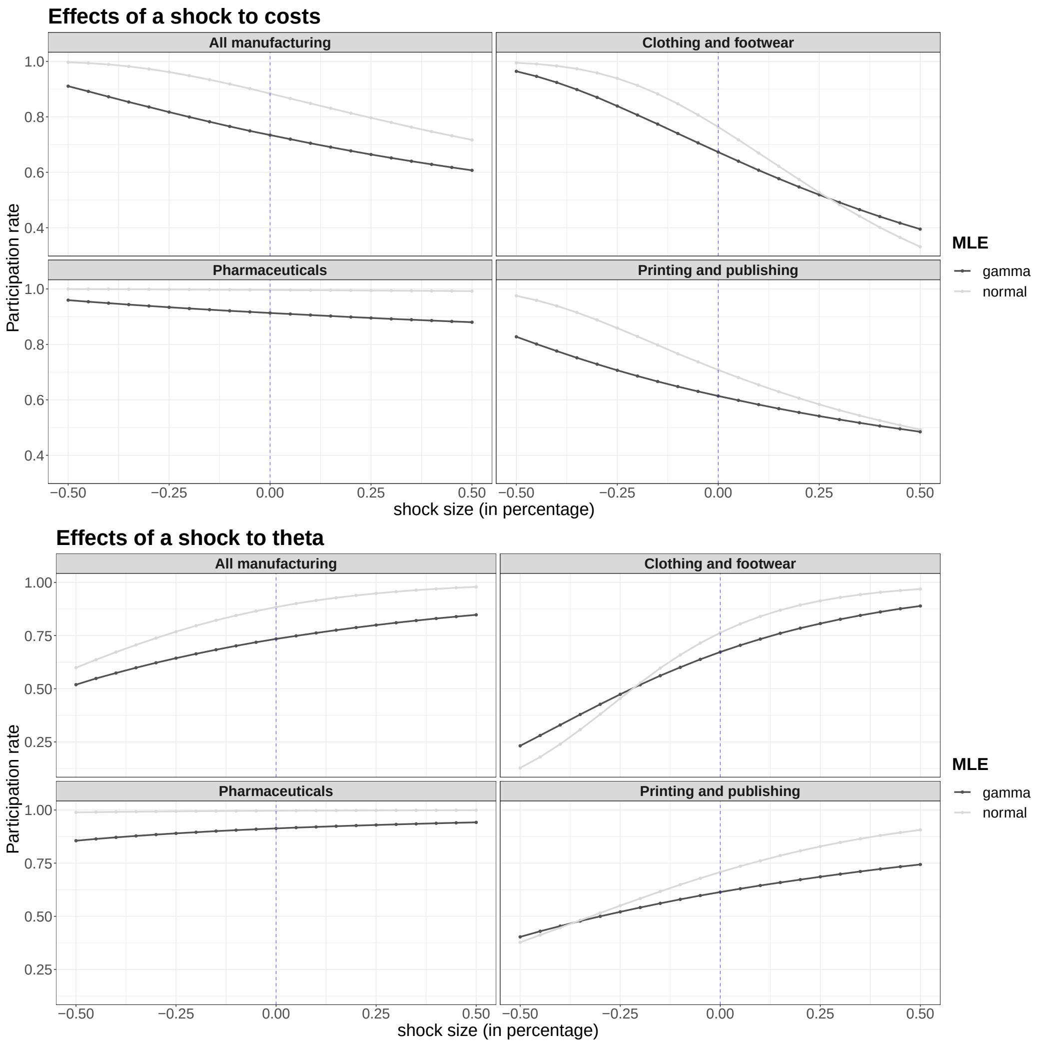

Figure 9: *Effects of a cost-shock (top four panels) and of a theta-shock (bottom four panels) on participation rates for selected sectors. Results are averaged across* 200 *MonteCarlo simulations.*

In Figure 9 we display, for the two shocks and for the four selected sectors of interest, the estimated participation rate (on the y-axis) as a function of the shock size τ (on the x-axis). For

[16] For every value of $\tau \in \{-0.5, -0.45, \ldots, +0.45, +0.5\}$ we perform 200 MonteCarlo stochastic replications. Here the MonteCarlo averages are reported.

$\tau = 0\%$ the participation rate equates the one reported in Table 8. Intuitively, there is a negative slope in the *cost-shock* case and a positive slope in the *theta-shock* case. As the costs rise, with constant productivities, fewer firms export. Oppositely, when the efficiency improves, with constant costs, more firms export.

Once again, the most interesting result is the one referring to the different distributional shapes – the Normal vs. the Gamma. We find that a right-skewed asymmetric distribution of thresholds is more inelastic with respect to cost-shocks. As a matter of fact even if the cost distribution slightly shifts to the left (due to a government subsidy to export for all firms in the specific sector), the core of firms with very low costs will continue to export, but only the fringe firms on the right tail of the distribution will benefit, revert their export decision and start to export.[17] A large proportion of firms with costs located on the right-tail of the distribution will instead be unaffected by the policy. Thus policies aimed at increasing the export participation and non-targeted to the specific firms on the right tail of the cost distribution would have only limited effects. On the opposite side, an asymmetric distribution is relatively inelastic also to slight increases in the costs. But a large shock to the threshold distribution however (e.g. a large tariff imposed by a foreign economy on domestic firms), might strongly affect the decisions of many firms, reducing participation rates even by more than 10 percentage points. The general implications, even if they slightly vary across sectors, are very similar (with reverse sign) also when looking at the theta-shock.

All in all, the asymmetric distribution is less elastic than the symmetric one. This implies a higher resilient to small negative shocks, but also a lower sensitivity to positive shocks. Therefore, if thresholds are heterogeneous and asymmetrically distributed as our results suggest, the policy efforts should be local, targeted to specific firms rather than global and widespread. Policies shall indeed aim at affecting higher order moments of the cost distribution, rather than affecting only the centrality moment.

6 Conclusion

In this paper we propose a parametric Maximum Likelihood Estimation (MLE) approach for the discovery of the parameters characterizing threshold distributions. In the first stage we show that the problem can be derived by assuming a distribution of heterogeneous thresholds, conditional on the observation of the decision outcome as well as a critical variable of each individual. The approach comprises minimal requirements. We also use stochastic MonteCarlo simulations in order to study the reliability of our approach when the baseline assumptions are not satisfied, to define the boundaries of its application. In a second stage we provide an empirical application to the problem of export participation of a sample of French firms. On the one side this application allows us to extend the theoretical concept of export threshold from a single and homogeneous cut-off, to a distribution of heterogeneous cut-offs. On the other side it allows us to rationalize the puzzle of overlapping productivities between exporters and non-exporters. A horse-race study of the empirical application also allows us test for the empirical relevance of the heterogeneous thresholds theory and of the asymmetric shape of the

[17] The already exporting firms might eventually benefit by increasing the intensive margin, which is not studied in this paper.

threshold distributions. We indeed estimate that the four lowest order moments of the threshold distributions are powerful tools to predict participation rates. In a third stage, we employ the estimates from our empirical exercise to investigate upon the possible of effects of policy shocks on the export participation rate. In particular, we study the effects of an exchange rate shock and of a export subsidy shock, finding that asymmetric distributions are less sensitive to both positive and negative shocks. Overall, our results indicate that accounting for agents heterogeneity and for higher order moments allows one to gain new relevant information, which can also be used for policy purposes. If threshold are asymmetric, broad policies which focus on the central moments are indeed ineffective. Targeted policies, affecting only specific group of agents shall be employed.

Our work can be extended in at least two directions. First, by broadening the meaning of the θ-attribute toward a multivariate context. With respect to the export application, for example, one can think that thresholds are heterogeneous also within firm. By looking at the problem at a more fine grained scale, it seems natural to believe that export barriers are different with respect to different locations (for firms exporting in multiple location) or with respect to the export of different product (for firms producing multiple products). In this case one shall extend the work by looking at the multivariate Gamma or Normal distributions, and take into account the correlation existing between the two dimensions. Second, within the univariate case, one can aim at localizing individual firms over the whole distribution of thresholds. This is feasible by employing additional firm characteristics and structural econometric techniques. Such an exercise would be particularly useful for policy purposes, with the scope of optimizing the allocation of export subsidies to those firms which are very productive but decide not to export because their thresholds are extremely high.

References

Alogoskoufis, G. and A. Manning (1991). Tests of alternative wage employment bargaining models with an application to the UK aggregate labour market. *European Economic Review 35*(1), 23 – 37.

Bernard, A. B. and J. B. Jensen (2004). Why some firms export. *The Review of Economics and Statistics 86*(2), 561–569.

Cogan, J. F. (1981). Fixed costs and labor supply. *Econometrica 49*(4), 945–963.

Dixit, A. (1989). Entry and exit decisions under uncertainty. *Journal of Political Economy 97*(3), 620–638.

Eaton, J., S. Kortum, and F. Kramarz (2011, September). An anatomy of international trade: Evidence from french firms. *Econometrica 79*, 1453–1498.

Impullitti, G., A. A. Irarrazabal, and L. D. Opromolla (2013). A theory of entry into and exit from export markets. *Journal of International Economics 90*(1), 75 – 90.

Kau, P. and L. Hill (1972). A threshold model of purchasing decisions. *Journal of Marketing Research 9*(3), 264–270.

Mayer, T., M. J. Melitz, and G. I. P. Ottaviano (2014, February). Market size, competition, and the product mix of exporters. *American Economic Review 104*(2), 495–536.

McDonald, R. and D. Siegel (1986, 11). The Value of Waiting to Invest*. *The Quarterly Journal of Economics 101*(4), 707–727.

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

Melitz, M. J. and G. I. P. Ottaviano (2008, 01). Market size, trade, and productivity. *The Review of Economic Studies 75*(1), 295–316.

Pissarides, C. A. (1974). Risk, job search, and income distribution. *Journal of Political Economy 82*(6), 1255–1267.

Wei, Z. and H. Timmermans (2008). Cut-off models for the go-home decision of pedestrians in shopping streets. *Environment and Planning B: Planning and Design 35*(2), 248–260.

A Appendix

A.1 Algorithms for the MonteCarlo simulation exercises

Baseline MonteCarlo Settings

The MonteCarlo simulations are carried out as follows:

1. fix a sufficiently large number of agents N;
2. simulate the *true* θ-attribute data $\boldsymbol{\theta}^T$ from a known distribution g;
3. simulate the *true* threshold data $\boldsymbol{c}^T$ from a known distribution f;
4. let the agents compute their individual decision outcomes $\boldsymbol{\chi}$ according to equation 1;
5. using the available information to the social researcher (i.e. $\boldsymbol{\theta}$ and $\boldsymbol{\chi}$) estimate with the maximum likelihood the parameters $\hat{\Omega}$ that characterize the threshold distribution f;
6. repeat steps 2 to 5 a sufficient number of time M;
7. use the M estimates $\hat{\boldsymbol{\Omega}}$ to evaluate the goodness of the estimation.

MonteCarlo Settings - Testing Assumption A1

The MonteCarlo simulations are carried out as follows:

1. fix a sufficiently large number of agents N;
2. fix a set of probability density functions $\mathcal{F} = f_1, f_2, \ldots, f_K$;
3. simulate the *true* θ-attribute data $\boldsymbol{\theta}^T$ from a known distribution g;
4. simulate the *true* threshold data $\boldsymbol{c}^T$ from the known distribution f_k;
5. let the agents compute their individual decision outcomes $\boldsymbol{\chi}$ according to equation 1;
6. using the available information to the social researcher (i.e. $\boldsymbol{\theta}$ and $\boldsymbol{\chi}$) estimate with the maximum likelihood the parameters $\hat{\Omega}$ that characterize the threshold distribution f_k;
7. repeat steps 3 to 6 a sufficient number of time M;
8. use the M estimates $\hat{\boldsymbol{\Omega}}$ to evaluate the goodness of the estimation;
9. repeat steps 3 to 8 for all the probability density functions in $\mathcal{F}$, as defined at step 2;
10. evaluate and compare the goodness of all the density functions in $\mathcal{F}$.

MonteCarlo Settings - Testing Assumption A2

The MonteCarlo simulations are carried out as follows:

1. fix a sufficiently large number of agents N;
2. fix a vector of noise $\boldsymbol{\sigma} = \sigma_1, \sigma_2, \ldots, \sigma_K$;
3. simulate the *true* θ-attribute data $\boldsymbol{\theta}^T$ from a known distribution g;
   - generate also the noisy θ-attribute data $\boldsymbol{\theta}^\varepsilon = \boldsymbol{\theta}^T + \boldsymbol{\varepsilon}^\theta$;
4. simulate the *true* threshold data $\boldsymbol{c}^T$ from the known distribution f_k;
   - generate also the noisy threshold data $\boldsymbol{c}^\varepsilon = \boldsymbol{c}^T + \boldsymbol{\varepsilon}^c$;
5. let the agents compute their individual decision outcomes $\boldsymbol{\chi}$ according to equation 7;
6. using the information available to the social researcher estimate with the maximum likelihood the parameters $\hat{\Omega}$ that characterize the threshold distribution f_k;
7. repeat steps 3 to 6 a sufficient number of time M;
8. use the M estimates $\hat{\boldsymbol{\Omega}}$ to evaluate the goodness of the estimation;
9. repeat steps 3 to 8 for all the noise levels $\boldsymbol{\sigma}$, as defined at step 2;
10. evaluate and compare the goodness of all the values of $\boldsymbol{\sigma}$.

A.2 Distributions

Univariate Normal distribution

In the case of a Normal distribution, the probability density function is defined as:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \tag{10}$$

where μ and σ represent the average and the standard deviation, respectively. This distribution is typically denoted as $x \sim \mathcal{N}(\mu, \sigma^2)$. Integrating over the interval $(-\infty, \bar{x}]$ yields the cumulative density function:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} exp \int_{-\infty}^{\bar{x}} \left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) dx \tag{11}$$

Univariate Gamma distribution

In the case of a Gamma distribution, the probability density function is defined as:

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} exp\left(\frac{-x}{\beta}\right) \tag{12}$$

where $\alpha > 0$ and $\beta > 0$ represent the shape and scale parameters, respectively and $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} exp(-t)dt$ is the Gamma function. The mean and variance of such distribution are combination of shape and scale parameters and read respectively $\mu = \alpha\beta$ and $\sigma = \alpha\beta^2$. This distribution is typically denoted as $x \sim \Gamma(\alpha, \beta)$. Integrating over the interval $(0, \bar{x}]$ yields the cumulative density function:

$$F(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{\bar{x}} x^{\alpha-1} exp\left(\frac{-x}{\beta}\right) dx \tag{13}$$

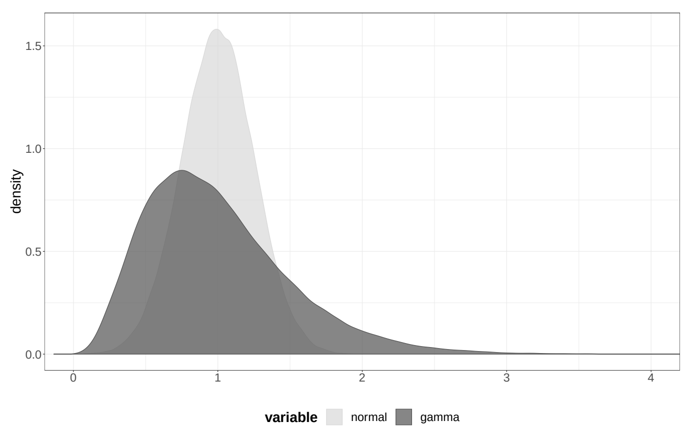

Figure 10: *Examples of Normal and Gamma distributions with equal mean and variance. In particular:* $\mu = 1, \sigma^2 = 0.25$.

A.3 Details of the empirical application and robustness checks

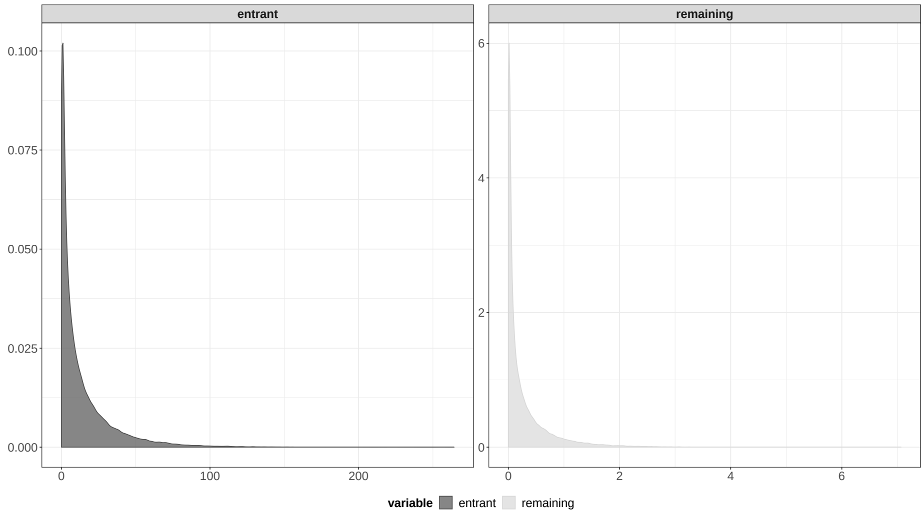

Figure 11: *Comparison of entry and remaining threshold distribution according to the Gamma MLE estimation for the aggregate manufacturing sector. The (α, β) parameters employed for the distributions are the ones estimated according to the first row of the top and bottom panels of Table 6. The scales are different to display the vast difference in the support of export thresholds for entrant and remaining firms.*

INCLUDE HERE:

- explanation of how we calculated both the Total Factor Productivity and the Apparent Labour Productivity (ALP)
- results of the estimation for participation/entry/remaining with ALP
- results of the horse race for participation/entry/remaining with ALP