# Online Appendix to "Unobserved Heterogeneity in the Productivity Distribution and Gains From Trade"

Ruben Dewitte, Michel Dumont, Glenn Rayp, Peter Willemé*

26th May 2020

*Dewitte: Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium, rubenl.dewitte@ugent.be, Dumont: Federal Planning Bureau and Ghent University, Rue Belliard - Belliardstraat 14-18 1040 Brussels, dm@plan.be , Rayp: Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium, glenn.rayp@ugent.be, Willemé: Federal Planning Bureau and Ghent University, Rue Belliard - Belliardstraat 14-18 1040 Brussels, pw@plan.be. We thank Pete Klenow, Luca David Opromolla, Gonzague Vannoorenberghe, Gerdie Everaert, Bruno Merlevede, Marijn Verschelde and participants at the 2018 ETSG conference in Warsaw, the 2019 ETSG conference in Bern, the DEGIT XIV conference in Odense and the 2019 workshop on "Firm Heterogeneity in Technical Change" in Ghent for their helpful comments and suggestions.

# Appendix A Additional Figures and table

## A.1 Figures

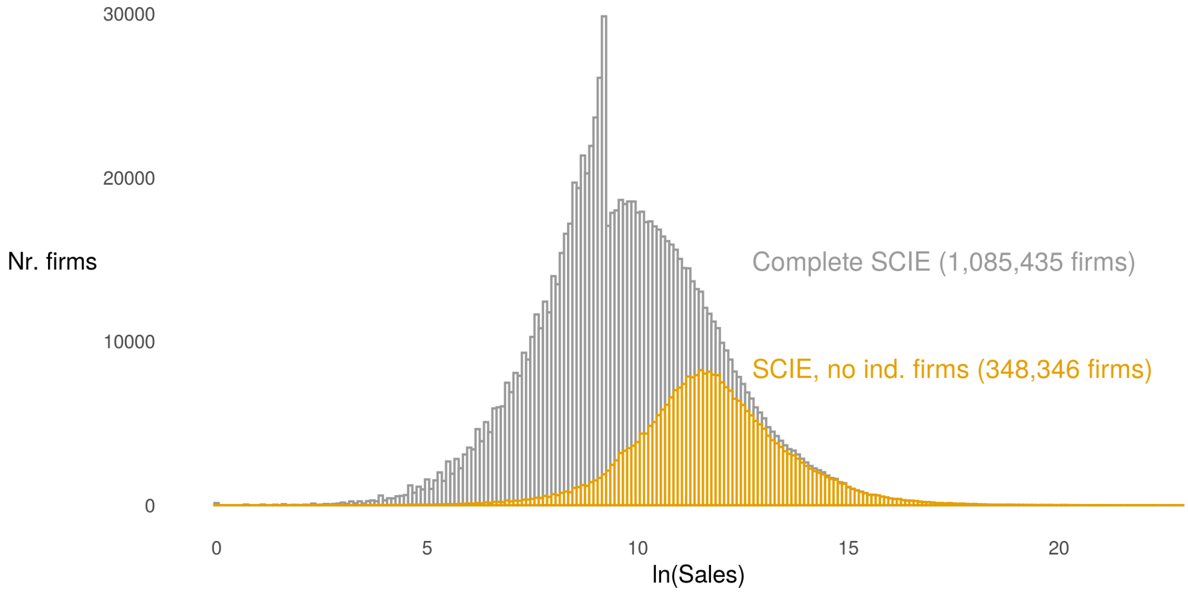


Figure 1: Density comparison of the SCIE dataset with and without individual companies.

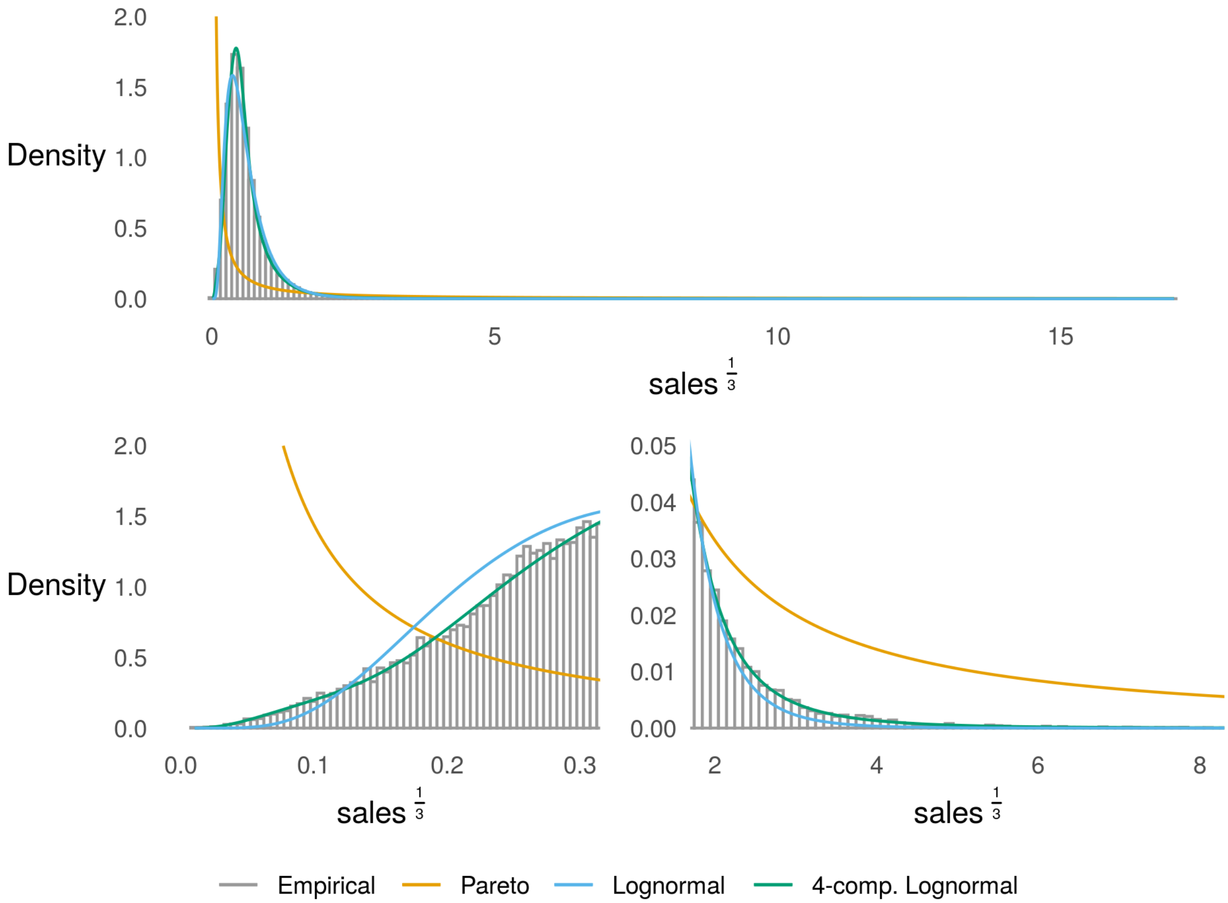


Figure 2: Empirical probability density function of Portuguese firm productivity in 2006 (upper panel) with fitted Pareto and (4-component) Lognormal densities. The lower left and right panels focus in on the left and right tail respectively.

**Notes:** Productivity is measured as domestic sales (relative to the mean) to the power of $1/(\sigma - 1)$ with $\sigma$, the elasticity of substitution between varieties, set to four. Distributions are fitted using maximum likelihood methods (cf. infra) to the complete dataset.

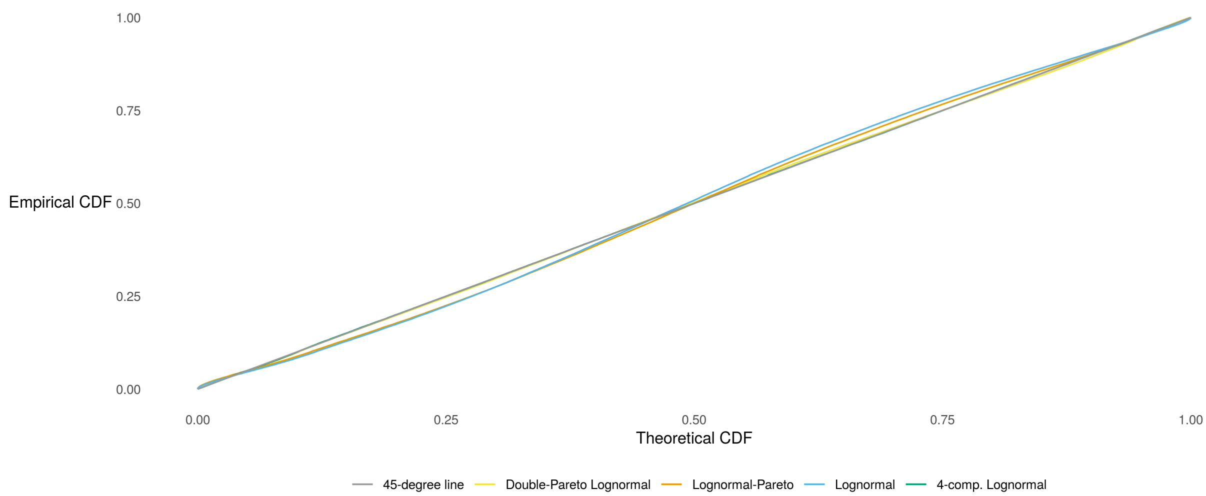


Figure 3: Probability-Probability plot for the Double-Pareto Lognormal, Lognormal-Pareto, Lognormal and 4-component Lognormal over the complete range of domestic sales in Portugal, 2006.

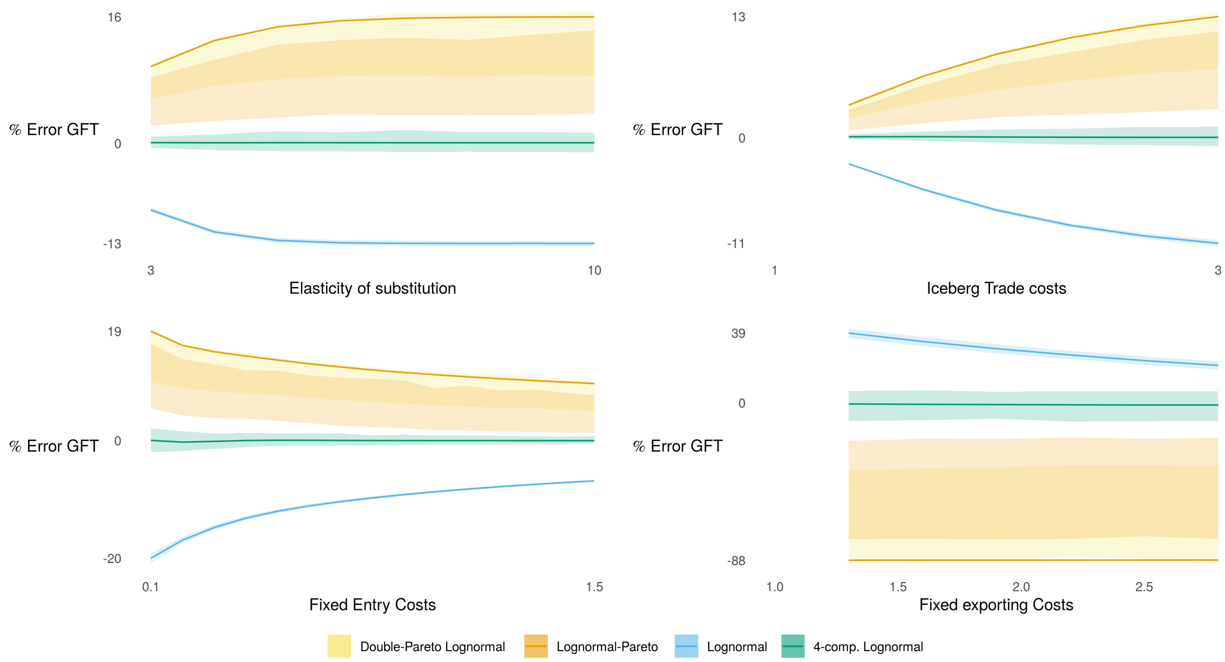


Figure 4: Percentage errors in parametric GFT calculations relative to the empirical benchmark for different values of the elasticity of substitution (left upper panel) and different fixed entry costs (left bottom panel) for a reduction in variable trade costs $\left(\tau^{ij} = 3 \rightarrow (\tau^{ij})' = 1\right)$. The right upper panel displays percentage errors in parametric GFT for different starting values of the iceberg trade costs $\left(\tau^{ij} \in [1;3] \rightarrow (\tau^{ij})' = 1\right)$. The bottom left panel showcases the error in parametric GFT for a reduction in fixed exporting costs with different starting values $\left(f^{ij} \in [1;3] \rightarrow (f^{ij})' = 1\right)$.

**Notes:** Full lines represent the parametric population GFT, while shaded areas delineate the 5th and 95th quantile of the parametric bootstrapped (999 replications) finite sample GFT. The Double-Pareto Lognormal has no finite population GFT value.

## A.2 Tables

Table 1: Overview of all distributions considered.

| Distribution | Abbreviation | Support | Parameters | Change in parameters from power transformation $ax^b$ |
|---|---|---|---|---|
| Pareto | P | $[x_{min}, \infty[$ | $k, x_{min}$ | $kb, \left(\frac{x_{min}}{a}\right)^{\frac{1}{b}}$ |
| Inverse Pareto | IP | $[0, x_{max}]$ | $k, x_{max}$ | $kb, \left(\frac{x_{max}}{a}\right)^{\frac{1}{b}}$ |
| Lognormal | LN | $[0, \infty[$ | $\mu, Var$ | $\frac{\mu - lna}{b}, \frac{Var}{b}$ |
| Weibull | W | $[0, \infty[$ | $k, s$ | $bk, \left(\frac{s}{a}\right)^{\frac{1}{b}}$ |
| Exponential | Exp | $[0, \infty[$ | $s$ | $W\left(b, \left(\frac{s}{a}\right)^{\frac{1}{b}}\right)$ |
| Burr | B | $[0, \infty[$ | $k, c, s$ | $k, bc, \left(\frac{s}{a}\right)^{\frac{1}{b}}$ |
| Fréchet | F | $[0, \infty]$ | $k, s$ | $bk, \left(\frac{s}{a}\right)^{\frac{1}{b}}$ |
| Generalized Gamma | GG | $[0, \infty[$ | $k, c, s$ | $bk, bc, \left(\frac{s}{a}\right)^{\frac{1}{b}}$ |
| Gamma | G | $[0, \infty[$ | $k, s$ | $GG\left(bk, b, \left(\frac{s}{a}\right)^{\frac{1}{b}}\right)$ |
| Finite Mixture Model | FMM | See ind. comp. | $\boldsymbol{\Psi}$ | See ind. comp. |
| Piecewise composite | PC | See ind. comp. | $\boldsymbol{\theta}$ | See ind. comp. |
| Double-Pareto Lognormal | DPLN | $[0, \infty[$ | $k_1, \mu, Var, k_2$ | $\frac{k_1}{b}, b\mu + log(a), Var, \frac{k_2}{b}$ |
| Left-Pareto Lognormal | LPLN | $[0, \infty[$ | $k_1, \mu, Var$ | $\frac{k_1}{b}, b\mu + log(a), Var$ |
| Right-Pareto Lognormal | RPLN | $[0, \infty[$ | $\mu, Var, k_2$ | $b\mu + log(a), Var, \frac{k_2}{b}$ |

Table 2: Overview of the probability and cumulative density functions of single distributions considered.

| Distribution | PDF | CDF |
|---|---|---|
| P | $\frac{kx_{min}^{k}}{x^{k+1}}$ | $1-\left(\frac{x_{min}}{x}\right)^{k}$ |
| IP | $\frac{kx_{max}^{-k}}{x^{-k+1}}$ | $1-\left(\frac{x_{max}}{x}\right)^{-k}$ |
| LN | $\frac{1}{xVar\sqrt{2\pi}}e^{-(lnx-\mu)^2/2Var^2}$ | $\Phi\left(\frac{lnx-\mu}{Var}\right)$ |
| W | $\frac{k}{s}\left(\frac{x}{s}\right)^{k-1}e^{-\left(\frac{x}{s}\right)^{k}}$ | $1-e^{-\left(\frac{x}{s}\right)^{k}}$ |
| Exp | $\frac{1}{s}e^{-\frac{x}{s}}$ | $1-e^{-\frac{x}{s}}$ |
| B | $\frac{\frac{kc}{s}\left(\frac{x}{s}\right)^{c-1}}{\left(1+\left(\frac{x}{s}\right)^{c}\right)^{k+1}}$ | $1-\frac{1}{\left(1+\left(\frac{x}{s}\right)^{c}\right)^{k}}$ |
| F | $\frac{k}{s}\left(\frac{x}{s}\right)^{-1-k}e^{-\left(\frac{x}{s}\right)^{-k}}$ | $e^{-\left(\frac{x}{s}\right)^{-k}}$ |
| GG[a] | $\frac{c}{s^k\Gamma(\frac{k}{c})}x^{k-1}e^{-\left(\frac{x}{s}\right)^{c}}$ | $\frac{1}{\Gamma(\frac{k}{c})}\gamma(\frac{k}{c},\left(\frac{x}{s}\right)^{c})$ |
| G[a] | $\frac{1}{s^k\Gamma(k)}x^{k-1}e^{-\frac{x}{s}}$ | $\frac{1}{\Gamma(k)}\gamma(k,\frac{x}{s})$ |

**Notes**: [a]$\Gamma(x)$ stands for the Gamma function, while $\gamma(s,x)$ stands for the lower incomplete Gamma function with upper bound $x$.

Table 3: Overview of the probability and cumulative density functions of combined distributions considered.

| **Distribution** | **PDF** | **CDF** |
| --- | --- | --- |
| FMM | $\sum_{i=1}^{I} \pi_i m_i(x\vert\boldsymbol{\theta}_i)$ | $\sum_{i=1}^{I} \pi_i M(x\vert\boldsymbol{\theta}_i)$ |
| PC[a] | $\begin{cases} \frac{\alpha_1}{1+\alpha_1+\alpha_2} m_1^*(x\vert\boldsymbol{\theta}_1) & \text{if} \quad 0 < x \leq c_1 \\ \frac{1}{1+\alpha_1+\alpha_2} m_2^*(x\vert\boldsymbol{\theta}_2) & \text{if} \quad c_1 < x \leq c_2 \\ \frac{\alpha_2}{1+\alpha_1+\alpha_2} m_3^*(x\vert\boldsymbol{\theta}_3) & \text{if} \quad c_2 < x < \infty \end{cases}$ | $\begin{cases} \frac{\alpha_1}{1+\alpha_1+\alpha_2} \frac{M_1(x\vert\boldsymbol{\theta}_1)}{M_1(c_1\vert\boldsymbol{\theta}_1)} & \text{if} \quad 0 < x \leq c_1 \\ \frac{\alpha_1}{1+\alpha_1+\alpha_2} + \frac{1}{1+\alpha_1+\alpha_2} \frac{M_2(x\vert\boldsymbol{\theta}_2)-M_2(c_1\vert\boldsymbol{\theta}_2)}{M_2(c_2\vert\boldsymbol{\theta}_2)-M_2(c_1\vert\boldsymbol{\theta}_2)} & \text{if} \quad c_1 < x \leq c_2 \\ \frac{1+\alpha_1}{1+\alpha_1+\alpha_2} + \frac{\alpha_2}{1+\alpha_1+\alpha_2} \frac{M_3(x\vert\boldsymbol{\theta}_3)-M_3(c_2\vert\boldsymbol{\theta}_3)}{1-M_3(c_2\vert\boldsymbol{\theta}_3)} & \text{if} \quad c_2 < x < \infty \end{cases}$ |
| DPLN[b] | $\frac{k_2 k_1}{k_2+k_1}\left[x^{-k_2-1} e^{k_2\mu+\frac{k_2^2 Var^2}{2}} \Phi\left(\frac{lnx-\mu-k_2 Var^2}{Var}\right) + x^{k_1-1} e^{-k_1\mu+\frac{k_1^2 Var^2}{2}} \Phi^c\left(\frac{lnx-\mu+k_1 Var^2}{Var}\right)\right]$ | $\Phi\left(\frac{lnx-\mu}{Var}\right) - \frac{1}{k_2+k_1}\left[k_1 x^{-k_2} e^{k_2\mu+\frac{k_2^2 Var^2}{2}} \Phi\left(\frac{lnx-\mu-k_2 Var^2}{Var}\right) - k_2 x^{k_1} e^{-k_1\mu+\frac{k_1^2 Var^2}{2}} \Phi^c\left(\frac{lnx-\mu+k_1 Var^2}{Var}\right)\right]$ |
| LPLN[b] | $k_1 x^{k_1-1} e^{-k_1\mu+\frac{k_1^2 Var^2}{2}} \Phi^c\left(\frac{lnx-\mu+k_1 Var^2}{Var}\right)$ | $\Phi\left(\frac{lnx-\mu}{Var}\right) - x^{k_1} e^{-k_1\mu+\frac{k_1^2 Var^2}{2}} \Phi^c\left(\frac{lnx-\mu+k_1 Var^2}{Var}\right)$ |
| RPLN[b] | $k_2 x^{-k_2-1} e^{k_2\mu+\frac{k_2^2 Var^2}{2}} \Phi\left(\frac{lnx-\mu-k_2 Var^2}{Var}\right)$ | $\Phi\left(\frac{lnx-\mu}{Var}\right) - x^{-k_2} e^{k_2\mu+\frac{k_2^2 Var^2}{2}} \Phi\left(\frac{lnx-\mu-k_2 Var^2}{Var}\right)$ |

**Notes:** [a] $\forall i \in I : m_i^*(x) = \frac{m_i(x)}{\int_{c_{i-1}}^{c_i} m_i(x)dx}$, [b] $\Phi$ and $\Phi^c$ stand for the standard normal and complementary standard normal cdfs.

Table 4: Expression of the $y$-bounded $r$th moment ($\mu_y^r$) for the single distributions considered.

| **Distribution** | $\mu_y^r$ | **Additional parameter restrictions**[a] |
|---|---|---|
| P | $-\left(y\right)^{r-k}\,\frac{k\omega_{min}^k}{r-k}$ | $k > r$ |
| IP | $k\omega_{max}^{-k}\frac{\left(\omega_{max}\right)^{r+k}-\left(y\right)^{r+k}}{r+k}$ | - |
| LN | $e^{\frac{r\left(rVar^2+2\mu\right)}{2}}\left[1-\Phi\left(\frac{lny-\left(rVar^2+\mu\right)}{Var}\right)\right]$ | - |
| W[c] | $s^{\sigma_s-1}\Gamma\left(\frac{\sigma_s-1}{k}+1,\left(\frac{y}{s}\right)^k\right)$ | - |
| Exp[c] | $s^{\sigma_s-1}\Gamma\left(\sigma_s+1,\frac{y}{s}\right)$ | - |
| B[b] | $s^r k\left[\boldsymbol{B}\left(\frac{r}{c}+1,k-\frac{r}{c}\right)-\boldsymbol{B}\left(\frac{\left(\frac{y}{s}\right)^c}{1+\left(\frac{y}{s}\right)^c};\frac{r}{c}+1,k-\frac{r}{c}\right)\right]$ | $c > r,\ kc > r$ |
| F[c] | $s^{\sigma_s-1}\left[1-\Gamma\left(1-\frac{\sigma_s-1}{k},\left(\frac{y}{s}\right)^{-k}\right)\right]$ | $k > r$ |
| GG[c] | $\frac{s^{\sigma_s-1}}{\Gamma\left(\frac{k}{c}\right)}\Gamma\left(\frac{\sigma_s-1+k}{c},\left(\frac{y}{s}\right)^c\right)$ | - |
| G[c] | $\frac{s^{\sigma_s-1}}{\Gamma(k)}\Gamma\left(\sigma_s-1+k,\frac{y}{s}\right)$ | - |

**Notes:** [a] Additional parameter restrictions represent parameter restrictions needed to keep the statistic finite. [b] $\boldsymbol{B}(a,b)$ stands for the beta function, while $\boldsymbol{B}(x,a,b)$ stands for the lower incomplete beta function with upper bound $x$. [c] $\Gamma(x)$ stands for the Gamma function, while $\Gamma(s,x)$ stands for the upper incomplete Gamma function with lower bound $x$.

Table 5: Expression of the $y$-bounded $r$th moment ($\mu_y^r$) for the combined considered.

| **Distribution** | $\mu_y^r$ | **Additional parameter restrictions**[a] |
|---|---|---|
| FMM | $\sum_{i=1}^{I} \pi_i (\mu_i)_y^r$ | See ind. comp. |
| PC | $\begin{cases} \frac{\alpha_1}{1+\alpha_1+\alpha_2}\frac{(\mu_1)_y^r-(\mu_1)_{c_1}^r}{M_1(c_1)} + \frac{1}{1+\alpha_1+\alpha_2}\frac{(\mu_2)_{c_1}^r-(\mu_2)_{c_2}^r}{M_2(c_2)-M_2(c_1)} + \frac{\alpha_2}{1+\alpha_1+\alpha_2}\frac{(\mu_3)_y^r}{1-M_3(c_2)} & \text{if} \quad 0 < y \leq c_2 \\ \frac{1}{1+\alpha_1+\alpha_2}\frac{(\mu_2)_y^r-(\mu_2)_{c_2}^r}{M_2(c_2)-M_2(c_1)} + \frac{\alpha_2}{1+\alpha_1+\alpha_2}\frac{(\mu_3)_{c_2}^r}{1-M_3(c_2)} & \text{if} \quad c_1 < y \leq c_2 \\ \frac{\alpha_2}{1+\alpha_1+\alpha_2}\frac{(\mu_3)_y^r}{1-M_3(c_2)} & \text{if} \quad c_2 < y < \infty \end{cases}$ | See ind. comp. |
| DPLN | $\begin{aligned} &-\frac{k_2k_1}{k_2+k_1}e^{k_2\mu+\frac{k_2^2Var^2}{2}}\frac{y^{\sigma_s-k_2-1}}{\sigma_s-k_2-1}\Phi\left(\frac{lny-\mu-k_2Var^2}{Var}\right) \\ &-\frac{k_2k_1}{k_2+k_1}\frac{1}{r-k_2}e^{\frac{r^2Var^2+2\mu r}{2}}\Phi^c\left(\frac{lny-rVar^2-\mu}{Var}\right) \\ &-\frac{k_2k_1}{k_2+k_1}e^{-k_1\mu+\frac{k_1^2Var^2}{2}}\frac{y^{\sigma_s+k_1-1}}{\sigma_s+k_1-1}\Phi^c\left(\frac{lny-\mu+k_1Var^2}{Var}\right) \\ &+\frac{k_2k_1}{k_2+k_1}\frac{1}{r+k_1}e^{\frac{r^2Var^2+2\mu r}{2}}\Phi^c\left(\frac{lny-rVar^2-\mu}{Var}\right) \end{aligned}$ | $k_2 > r$ |
| LPLN | $\begin{aligned} &-k_1e^{-k_1\mu+\frac{k_1^2Var^2}{2}}\frac{y^{\sigma_s+k_1-1}}{\sigma_s+k_1-1}\Phi^c\left(\frac{lny-\mu+k_1Var^2}{Var}\right) \\ &+\frac{k_1}{r+k_1}e^{\frac{r^2Var^2+2\mu r}{2}}\Phi^c\left(\frac{lny-rVar^2+\mu}{Var}\right) \end{aligned}$ | - |
| RPLN | $\begin{aligned} &-k_2e^{k_2\mu+\frac{k_2^2Var^2}{2}}\frac{y^{\sigma_s-k_2-1}}{\sigma_s-k_2-1}\Phi\left(\frac{lny-\mu-k_2Var^2}{Var}\right) \\ &-\frac{k_2}{r-k_2}e^{\frac{r^2Var^2+2\mu r}{2}}\Phi^c\left(\frac{lny-rVar^2+\mu}{Var}\right) \end{aligned}$ | $k_2 > r$ |

**Notes:** [a] Additional parameter restrictions represent parameter restrictions needed to keep the statistic finite.

Table 6: Coverage ratio of SCIE vs OECD SDBS database.

| | Number of Enterprises | | | | | | Total Employment | | | | | | Turnover | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NACE Rev.2 | 1-9 | 10-19 | 20-49 | 50-249 | > 250 | Total | 1-9 | 10-19 | 20-49 | 50-249 | > 250 | Total | 1-9 | 10-19 | 20-49 | 50-249 | > 250 | Total |
| 13 | 100 | | | | 100 | 100 | | | | | | | | | | | | |
| 14 | 100 | 100 | 100 | 100 | | 100 | | 100 | 100 | | | | | 100 | 100 | | | |
| 15 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 | | | 100 |
| 16 | | | | 100 | 100 | 100 | | | | | | 100 | | | | | | 100 |
| 17 | 100 | 100 | 100 | 100 | 100 | 100 | | | | 100 | 100 | 100 | | | | 100 | 100 | 100 |
| 18 | 100 | 100 | 100 | 100 | 100 | 100 | | | | 100 | 100 | 100 | | | | 100 | 100 | 100 |
| 19 | 100 | 100 | 100 | 100 | 100 | 100 | | 100 | 100 | 100 | | | | 100 | 100 | 100 | | |
| 20 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 21 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 22 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 23 | | | | | 100 | 100 | | | | | | | | | | | | |
| 24 | 100 | 100 | 100 | 100 | 100 | 100 | | | | 100 | 100 | | | | | 100 | 100 | |
| 25 | 100 | 100 | 100 | 100 | 100 | 100 | | | | 100 | 100 | 100 | | | | 100 | 100 | 100 |
| 26 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 27 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 28 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 29 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 |
| 30 | 100 | 100 | 100 | 100 | 100 | 100 | | | | | | | | | | | | |
| 31 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 |
| 32 | 100 | 100 | 100 | 100 | 100 | 100 | | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 | |
| 33 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | | | | 100 | 100 | 100 | | | |
| 34 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | | | | 100 | 100 | 100 | |
| 35 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | | | | 100 | 100 | 100 | |
| 36 | 100 | 100 | 100 | 100 | 100 | 100 | | 100 | | 100 | 100 | 100 | | 100 | | 100 | 100 | 100 |
| 37 | 100 | 100 | 100 | 100 | | 100 | | | | 100 | | 100 | | | | 100 | | 100 |
| 40 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 41 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 | | | 100 |
| 45 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 50 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 51 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 52 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 55 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 60 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 61 | 100 | 100 | 100 | 100 | 100 | 100 | | 100 | | 100 | | 100 | | 100 | | 100 | | 100 |
| 62 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | | 100 | | | 100 | 100 | | 100 |
| 63 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 64 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 |
| 70 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 | | | 100 |
| 71 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | | | 100 | | | 100 | | | 100 |
| 72 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 | 100 | | | 100 | 100 | 100 |
| 73 | 100 | 100 | 100 | 100 | | 100 | | | | | | 100 | | | | | | 100 |
| 74 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

**Notes:** Each cell corresponds to the ratio of our dataset compared to the data from the OECD structural SDBS database for the year 2006. Size classes are based on total employment. Empty cells and absent industries are due to missing information from SBDS, even though the data is available in our SCIE database.

Table 7: Distribution fits to Portuguese domestic sales in 2006.

| | | Goodness of fit | | | | Information Criteria | | |
|---|---|---|---|---|---|---|---|---|
| Distribution | Parms. | $T_a^0$ | $S_b^0$ | $T_a^1$ | $S_b^1$ | Loglike | $R_{AIC}$ | $R_{BIC}$ |
| 5-comp. Lognormal | 14 | 0.18 | 0.11 | 3.08 | 2.37 | 12,776 | 1 | $3^{+++}$ |
| | | (0.10;0.25) | (0.08;0.32) | (2.03;9.73) | (0.82;26.24) | | | |
| 4-comp. Lognormal | 11 | 0.19 | 0.11 | 2.78 | 0.13 | 12,770 | 2 | 2 |
| | | (0.09;0.25) | (0.08;0.32) | (2.07;9.23) | (0.83;24.67) | | | |
| 5-comp. Burr | 19 | 0.19 | 0.12 | - | - | 12,767 | 3 | $7^{+++}$ |
| | | (0.10;0.25) | (0.08;0.32) | (-;-) | (-;-) | | | |
| 4-comp. Burr | 15 | 0.24 | 0.14 | - | - | 12,754 | 4 | $6^{+++}$ |
| | | (0.10;0.25)* | (0.08;0.32) | (-;-) | (-;-) | | | |
| 3-comp. Burr | 11 | 0.25 | 0.17 | - | - | 12,748 | 6 | $4^{+++}$ |
| | | (0.09;0.25)* | (0.08;0.30) | (-;-) | (-;-) | | | |
| 2-comp. Burr | 7 | 0.20 | 0.20 | - | - | 12,745 | 5 | 1 |
| | | (0.09;0.25) | (0.08;0.32) | (-;-) | (-;-) | | | |
| 5-comp. Weibull | 14 | 0.25 | 0.14 | 6.96 | 11.95 | 12,731 | 7 | $8^{+++}$ |
| | | (0.10;0.25)** | (0.08;0.31) | (1.29;5.00)*** | (0.59;13.45)* | | | |
| 3-comp. Lognormal | 8 | 0.29 | 0.34 | 4.39 | 9.91 | 12,723 | 8 | $5^{+++}$ |
| | | (0.10;0.24)** | (0.09;0.32)** | (2.34;11.34) | (0.93;30.68) | | | |
| 5-comp. Gamma | 14 | 0.26 | 0.16 | 7.27 | 0.09 | 12,639 | 9 | $9^{+++}$ |
| | | (0.10;0.26)** | (0.09;0.33) | (1.29;5.11)*** | (0.44;14.23) | | | |
| Inv. Pareto-Burr | 4 | 0.51 | 0.61 | - | - | 12,561 | 10 | $10^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.33)*** | (-;-) | (-;-) | | | |
| Inv. Pareto-Burr-Pareto | 5 | 0.51 | 0.61 | - | - | 12,561 | 11 | $11^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.33)*** | (-;-) | (-;-) | | | |
| 5-comp. Exponential | 9 | 0.32 | 0.23 | 7.96 | 0.15 | 12,548 | 12 | $12^{+++}$ |
| | | (0.09;0.26)*** | (0.09;0.31) | (1.31;4.78)*** | (0.40;12.83) | | | |
| 4-comp. Weibull | 11 | 0.31 | 0.25 | 14.75 | 27.04 | 12,543 | 13 | $13^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.31) | (0.87;3.44)*** | (0.29;8.78)*** | | | |
| Burr-Pareto | 4 | 0.73 | 0.95 | - | - | 12,451 | 15 | $15^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.33)*** | (-;-) | (-;-) | | | |
| Burr | 3 | 0.73 | 0.95 | - | - | 12,451 | 14 | $14^{+++}$ |
| | | (0.10;0.24)*** | (0.08;0.31)*** | (-;-) | (-;-) | | | |
| Double-Pareto Lognormal | 4 | 0.66 | 0.80 | - | - | 12,429 | 16 | $16^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.33)*** | (-;-) | (-;-) | | | |
| 2-comp. Lognormal | 5 | 0.53 | 0.71 | 8.70 | 10.15 | 12,401 | 17 | $17^{+++}$ |
| | | (0.10;0.24)*** | (0.09;0.32)*** | (1.32;5.87)** | (0.54;16.11) | | | |
| Inv. Pareto-Lognormal-Pareto | 4 | 0.81 | 1.01 | - | - | 12,231 | 18 | $18^{+++}$ |
| | | (0.09;0.26)*** | (0.08;0.34)*** | (-;-) | (-;-) | | | |
| 4-comp. Gamma | 11 | 0.40 | 0.63 | 11.95 | 0.26 | 12,173 | 19 | $19^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.32)*** | (1.00;3.92)*** | (0.26;10.38) | | | |
| Inv. Pareto-Fréchet-Pareto | 4 | 1.11 | 1.48 | - | - | 11,953 | 20 | $20^{+++}$ |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | (0.09;0.25)*** | (0.08;0.33)*** | (-;-) | (-;-) | | | |
| 3-comp. Weibull | 8 | 0.69 | 0.92 | 20.31 | 39.45 | 11,855 | 21 | $21^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.31)*** | (0.73;2.60)*** | (0.23;6.78)*** | | | |
| 4-comp. Exponential | 7 | 0.57 | 0.89 | 13.91 | 0.36 | 11,801 | 22 | $22^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.32)*** | (0.95;3.61)*** | (0.34;9.44) | | | |
| Inv. Pareto-Weibull-Pareto | 4 | 1.60 | 2.00 | - | - | 11,338 | 24 | $24^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (-;-) | (-;-) | | | |
| Weibull-Pareto | 3 | 1.60 | 2.00 | - | - | 11,338 | 23 | $23^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (-;-) | (-;-) | | | |
| Inv. Pareto-Gamma-Pareto | 4 | 1.70 | 2.17 | - | - | 11,249 | 26 | $26^{+++}$ |
| | | (0.10;0.26)*** | (0.08;0.35)*** | (-;-) | (-;-) | | | |
| Gamma-Pareto | 3 | 1.70 | 2.17 | - | - | 11,249 | 25 | $25^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (-;-) | (-;-) | | | |
| Inv. Pareto-Exponential-Pareto | 3 | 1.97 | 2.71 | - | - | 11,044 | 27 | $27^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.33)*** | (-;-) | (-;-) | | | |
| Exponential-Pareto | 2 | 2.00 | 2.83 | - | - | 11,012 | 28 | $28^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (-;-) | (-;-) | | | |
| 3-comp. Gamma | 8 | 1.00 | 1.56 | 19.47 | 0.62 | 10,288 | 29 | $29^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.32)*** | (0.73;2.75)*** | (0.30;7.18) | | | |
| Inv. Pareto-Lognormal | 3 | 3.02 | 4.26 | 45.72 | 127.97 | 9,198 | 30 | $30^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.31)*** | (0.43;1.67)*** | (0.16;4.36)*** | | | |
| Lognormal-Pareto | 3 | 2.56 | 3.78 | 562.07 | 1683.18 | 8,721 | 31 | $31^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (169.39;451.06)*** | (371.67;1342.23)*** | | | |
| 3-comp. Exponential | 5 | 1.64 | 2.60 | 22.73 | 0.87 | 8,387 | 32 | $32^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (0.67;2.36)*** | (0.21;6.27) | | | |
| Left-Pareto Lognormal | 3 | 3.23 | 4.91 | 46.00 | 127.70 | 8,059 | 33 | $33^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.32)*** | (0.41;1.58)*** | (0.18;4.05)*** | | | |
| Right-Pareto Lognormal | 3 | 2.82 | 4.38 | 19.27 | 49.88 | 8,028 | 34 | $34^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (3.10;11.90)** | (1.23;32.23)** | | | |
| Lognormal | 2 | 2.93 | 5.03 | 41.38 | 113.04 | 7,372 | 35 | $35^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.33)*** | (0.47;1.84)*** | (0.17;4.76)*** | | | |
| 2-comp. Weibull | 5 | 2.10 | 3.19 | 35.20 | 72.85 | 6,442 | 36 | $36^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.31)*** | (0.42;1.50)*** | (0.16;3.80)*** | | | |
| 2-comp. Fréchet | 5 | 6.92 | 10.64 | - | - | -3,041 | 37 | $37^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (-;-) | (-;-) | | | |
| 5-comp. Fréchet | 14 | 6.96 | 10.63 | - | - | -3,045 | 40 | $40^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.31)*** | (-;-) | (-;-) | | | |
| 3-comp. Fréchet | 8 | 6.96 | 10.63 | - | - | -3,046 | 38 | $38^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (-;-) | (-;-) | | | |
| 4-comp. Fréchet | 11 | 6.98 | 10.63 | - | - | -3,047 | 39 | $39^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.32)*** | (-;-) | (-;-) | | | |
| 2-comp. Gamma | 5 | 4.00 | 5.93 | 31.79 | 2.24 | -3,381 | 41 | $41^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | (0.45;1.67)*** | (0.14;4.45) | | | |
| 2-comp. Exponential | 3 | 7.06 | 11.51 | 37.63 | 3.23 | -18,112 | 42 | $42^{+++}$ |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | (0.10;0.25)*** | (0.08;0.33)*** | (0.38;1.40)*** | (0.14;3.52)* | | | |
| Inv. Pareto-Weibull | 3 | 9.18 | 16.52 | 54.06 | 123.38 | -29,711 | 43 | $44^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.31)*** | (0.26;0.92)*** | (0.11;2.22)*** | | | |
| Weibull | 2 | 9.18 | 16.51 | 54.06 | 123.40 | -29,713 | 44 | $43^{+++}$ |
| | | (0.09;0.25)*** | (0.09;0.32)*** | (0.25;0.90)*** | (0.15;2.20)*** | | | |
| Fréchet | 2 | 8.91 | 16.72 | - | - | -32,908 | 45 | $45^{+++}$ |
| | | (0.10;0.26)*** | (0.08;0.33)*** | (-;-) | (-;-) | | | |
| Fréchet-Pareto | 3 | 8.91 | 16.72 | - | - | -32,908 | 46 | $46.5^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.31)*** | (-;-) | (-;-) | | | |
| Inv. Pareto-Fréchet | 3 | 8.91 | 16.72 | - | - | -32,908 | 46 | $46.5^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.33)*** | (-;-) | (-;-) | | | |
| Inv. Pareto-Gamma | 3 | 20.93 | 32.98 | 50.26 | 9.56 | -104,785 | 48 | $48^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.33)*** | (0.22;0.76)*** | (0.13;1.81)*** | | | |
| Gamma | 2 | 20.98 | 33.03 | 50.29 | 9.58 | -104,878 | 49 | $49^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.32)*** | (0.22;0.76)*** | (0.11;1.71)*** | | | |
| Exponential | 1 | 44.64 | 79.71 | 60.73 | 16.76 | -299,935 | 50 | $50^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.33)*** | (0.15;0.49)*** | (0.11;1.08)*** | | | |
| Inv. Pareto-Exponential | 2 | 44.64 | 79.71 | 60.73 | 16.76 | -299,935 | 51 | $51^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.32)*** | (0.15;0.51)*** | (0.11;1.12)*** | | | |
| Pareto | 2 | 48.34 | 68.18 | - | - | -436,227 | 52 | $52^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.33)*** | (-;-) | (-;-) | | | |

**Notes:** All distributions fitted using Maximum Likelihood.
Values between parentheses report the 5th and 95th quantile of the parametric bootstrapped test statistic with 999 replications. ***, **, * indicate significance of this test at 1%, 5% and 10% respectively.
$^{+++}$, $^{++}$, $^{+}$ indicates the difference between this distribution's BIC and the first-ranked distribution in terms of BIC ($\Delta BIC$) providing strong evidence in favour of the first-ranked distribution ($\Delta BIC > 10$), moderate evidence ($6 < \Delta BIC \leq 10$) and weak evidence ($2 < \Delta BIC \leq 6$) respectively.
$_a$ Values multiplied by 100 for expositional purpose, $_b$ Values divided by 1,000 for expositional purpose.

Table 8: Distribution fits to domestic sales of the Portuguese manufacturing sector in 2006.

| | | Goodness of fit | | Information Criteria | | |
|---|---|---|---|---|---|---|
| Distribution | Parms. | $T_a^0$ | $S_b^0$ | Loglike | $R_{AIC}$ | $R_{BIC}$ |
| 5-comp. Burr | 19 | 0.24<br>(0.25;0.66) | 0.02<br>(0.03;0.12) | -2,095 | 7 | $10^{+++}$ |
| 5-comp. Lognormal | 14 | 0.28<br>(0.25;0.67) | 0.03<br>(0.03;0.12) | -2,095 | 3 | $5^{+++}$ |
| 4-comp. Burr | 15 | 0.24<br>(0.25;0.67) | 0.02<br>(0.03;0.12) | -2,096 | 5 | $6^{+++}$ |
| 3-comp. Burr | 11 | 0.23<br>(0.25;0.67) | 0.03<br>(0.03;0.12) | -2,099 | 4 | $3^{+++}$ |
| 2-comp. Burr | 7 | 0.22<br>(0.25;0.66) | 0.02<br>(0.03;0.12) | -2,099 | 1 | 1 |
| 3-comp. Lognormal | 8 | 0.28<br>(0.25;0.69) | 0.02<br>(0.03;0.12) | -2,101 | 2 | $2^{+++}$ |
| 4-comp. Lognormal | 11 | 0.27<br>(0.25;0.67) | 0.02<br>(0.03;0.12) | -2,101 | 6 | $4^{+++}$ |
| 5-comp. Weibull | 14 | 0.34<br>(0.26;0.65) | 0.03<br>(0.03;0.11) | -2,104 | 8 | $7^{+++}$ |
| 5-comp. Gamma | 14 | 0.31<br>(0.26;0.66) | 0.04<br>(0.03;0.12) | -2,114 | 9 | $8^{+++}$ |
| 4-comp. Weibull | 11 | 0.40<br>(0.26;0.65) | 0.04<br>(0.03;0.11) | -2,131 | 10 | $9^{+++}$ |
| 5-comp. Exponential | 9 | 1.29<br>(0.25;0.66)*** | 0.15<br>(0.03;0.12)*** | -2,171 | 11 | $13^{+++}$ |
| 4-comp. Gamma | 11 | 0.50<br>(0.26;0.65) | 0.09<br>(0.03;0.12) | -2,178 | 12 | $15^{+++}$ |
| Inv. Pareto-Fréchet-Pareto | 4 | 0.65<br>(0.25;0.65)* | 0.09<br>(0.03;0.12) | -2,187 | 13 | $11^{+++}$ |
| Inv. Pareto-Burr | 4 | 0.88<br>(0.26;0.65)*** | 0.13<br>(0.03;0.12)** | -2,197 | 14 | $12^{+++}$ |
| Inv. Pareto-Burr-Pareto | 5 | 0.88<br>(0.25;0.69)*** | 0.13<br>(0.03;0.12)** | -2,197 | 15 | $14^{+++}$ |
| 3-comp. Weibull | 8 | 0.83<br>(0.25;0.66)*** | 0.14<br>(0.03;0.11)** | -2,222 | 16 | $17^{+++}$ |
| 2-comp. Lognormal | 5 | 0.73<br>(0.26;0.67)** | 0.11<br>(0.03;0.12)* | -2,232 | 17 | $16^{+++}$ |
| Double-Pareto Lognormal | 4 | 1.08<br>(0.26;0.66)*** | 0.17<br>(0.03;0.12)*** | -2,245 | 18 | $18^{+++}$ |
| 4-comp. Exponential | 7 | 1.18<br>(0.26;0.66)*** | 0.16<br>(0.03;0.12)*** | -2,251 | 19 | $20^{+++}$ |
| Inv. Pareto-Lognormal-Pareto | 4 | 1.27<br>(0.25;0.66)*** | 0.18<br>(0.03;0.11)*** | -2,263 | 20 | $19^{+++}$ |
| Burr-Pareto | 4 | 1.18<br>(0.26;0.67)*** | 0.25<br>(0.03;0.12)*** | -2,284 | 22 | $22^{+++}$ |
| Burr | 3 | 1.18<br>(0.26;0.67)*** | 0.25<br>(0.03;0.12)*** | -2,284 | 21 | $21^{+++}$ |
| Inv. Pareto-Gamma-Pareto | 4 | 1.65<br>(0.25;0.65)*** | 0.28<br>(0.03;0.12)*** | -2,346 | 23 | $24^{+++}$ |
| Inv. Pareto-Weibull-Pareto | 4 | 1.62<br>(0.25;0.68)*** | 0.28<br>(0.03;0.12)*** | -2,348 | 24 | $25^{+++}$ |
| Gamma-Pareto | 3 | 1.58<br>(0.25;0.68)*** | 0.27<br>(0.03;0.12)*** | -2,355 | 26 | $26^{+++}$ |
| Weibull-Pareto | 3 | 1.58<br>(0.27;0.67)*** | 0.27<br>(0.03;0.11)*** | -2,355 | 27 | $27^{+++}$ |
| Exponential-Pareto | 2 | 1.57<br>(0.26;0.68)*** | 0.27<br>(0.03;0.12)*** | -2,355 | 25 | $23^{+++}$ |
| Inv. Pareto-Exponential-Pareto | 3 | 1.57<br>(0.26;0.67)*** | 0.27<br>(0.03;0.12)*** | -2,355 | 28 | $28^{+++}$ |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3-comp. Gamma | 8 | 1.01 | 0.20 | -2,408 | 29 | $29^{+++}$ |
| | | (0.26;0.68)*** | (0.03;0.12)*** | | | |
| 3-comp. Exponential | 5 | 1.44 | 0.26 | -2,608 | 30 | $30^{+++}$ |
| | | (0.25;0.65)*** | (0.03;0.12)*** | | | |
| Inv. Pareto-Lognormal | 3 | 4.18 | 0.81 | -2,875 | 31 | $31^{+++}$ |
| | | (0.26;0.65)*** | (0.03;0.12)*** | | | |
| 2-comp. Weibull | 5 | 2.19 | 0.43 | -2,918 | 32 | $32^{+++}$ |
| | | (0.25;0.66)*** | (0.03;0.12)*** | | | |
| Lognormal-Pareto | 3 | 3.25 | 0.64 | -3,051 | 33 | $33^{+++}$ |
| | | (0.25;0.67)*** | (0.03;0.12)*** | | | |
| Left-Pareto Lognormal | 3 | 4.39 | 0.89 | -3,103 | 34 | $34^{+++}$ |
| | | (0.25;0.65)*** | (0.03;0.11)*** | | | |
| Right-Pareto Lognormal | 3 | 3.51 | 0.73 | -3,143 | 35 | $35^{+++}$ |
| | | (0.26;0.66)*** | (0.03;0.12)*** | | | |
| Lognormal | 2 | 3.96 | 0.88 | -3,250 | 36 | $36^{+++}$ |
| | | (0.25;0.65)*** | (0.03;0.11)*** | | | |
| 2-comp. Gamma | 5 | 3.35 | 0.71 | -4,108 | 37 | $37^{+++}$ |
| | | (0.25;0.65)*** | (0.03;0.12)*** | | | |
| 5-comp. Fréchet | 14 | 8.11 | 1.79 | -4,863 | 38 | $40^{+++}$ |
| | | (0.26;0.67)*** | (0.03;0.12)*** | | | |
| 4-comp. Fréchet | 11 | 8.35 | 1.80 | -4,870 | 39 | $39^{+++}$ |
| | | (0.25;0.66)*** | (0.03;0.12)*** | | | |
| 3-comp. Fréchet | 8 | 8.59 | 1.82 | -4,881 | 40 | $38^{+++}$ |
| | | (0.26;0.67)*** | (0.03;0.12)*** | | | |
| 2-comp. Fréchet | 5 | 9.55 | 1.92 | -4,955 | 41 | $41^{+++}$ |
| | | (0.25;0.67)*** | (0.03;0.12)*** | | | |
| 2-comp. Exponential | 3 | 5.75 | 1.20 | -5,550 | 42 | $42^{+++}$ |
| | | (0.25;0.67)*** | (0.03;0.12)*** | | | |
| Inv. Pareto-Weibull | 3 | 9.91 | 2.42 | -8,321 | 44 | $44^{+++}$ |
| | | (0.26;0.65)*** | (0.03;0.12)*** | | | |
| Weibull | 2 | 9.91 | 2.42 | -8,321 | 43 | $43^{+++}$ |
| | | (0.26;0.67)*** | (0.03;0.11)*** | | | |
| Inv. Pareto-Fréchet | 3 | 10.04 | 2.61 | -9,885 | 46 | $46^{+++}$ |
| | | (0.26;0.68)*** | (0.03;0.12)*** | | | |
| Fréchet-Pareto | 3 | 10.04 | 2.61 | -9,885 | 47 | $47^{+++}$ |
| | | (0.25;0.69)*** | (0.03;0.13)*** | | | |
| Fréchet | 2 | 10.04 | 2.61 | -9,885 | 45 | $45^{+++}$ |
| | | (0.25;0.65)*** | (0.03;0.11)*** | | | |
| Inv. Pareto-Gamma | 3 | 20.68 | 4.43 | -17,309 | 48 | $48^{+++}$ |
| | | (0.26;0.67)*** | (0.03;0.11)*** | | | |
| Gamma | 2 | 20.72 | 4.44 | -17,318 | 49 | $49^{+++}$ |
| | | (0.25;0.67)*** | (0.03;0.12)*** | | | |
| Exponential | 1 | 43.48 | 10.42 | -41,128 | 50 | $50^{+++}$ |
| | | (0.27;0.66)*** | (0.03;0.12)*** | | | |
| Inv. Pareto-Exponential | 2 | 43.48 | 10.42 | -41,128 | 51 | $51^{+++}$ |
| | | (0.26;0.65)*** | (0.03;0.11)*** | | | |
| Pareto | 2 | 49.14 | 9.43 | -66,043 | 52 | $52^{+++}$ |
| | | (0.26;0.65)*** | (0.03;0.11)*** | | | |

**Notes:** All distributions fitted using Maximum Likelihood.
Values between parentheses report the 5th and 95th quantile of the parametric bootstrapped test statistic with 999 replications. ***, **, * indicate significance of this test at 1%, 5% and 10% respectively.
$^{+++}$, $^{++}$, $^{+}$ indicates the difference between this distribution's BIC and the first-ranked distribution in terms of BIC ($\Delta BIC$) providing strong evidence in favour of the first-ranked distribution ($\Delta BIC > 10$), moderate evidence ($6 < \Delta BIC \leq 10$) and weak evidence ($2 < \Delta BIC \leq 6$) respectively.
$_a$ Values multiplied by 100 for expositional purpose, $_b$ Values divided by 1,000 for expositional purpose.

Table 9: Distribution fits to Portuguese domestic sales leaving out the first and last 1,000 observations in 2006.

| | | Goodness of fit | | Information Criteria | | |
|---|---|---|---|---|---|---|
| Distribution | Parms. | $T_a^0$ | $S_b^0$ | Loglike | $R_{AIC}$ | $R_{BIC}$ |
| 4-comp. Lognormal | 11 | 0.18 | 0.18 | 23,100 | 1 | 1 |
| | | (0.09;0.24) | (0.09;0.32) | | | |
| 5-comp. Lognormal | 14 | 0.21 | 0.20 | 23,093 | 2 | $2^{+++}$ |
| | | (0.09;0.24) | (0.08;0.30) | | | |
| 3-comp. Lognormal | 8 | 0.25 | 0.28 | 22,844 | 3 | $3^{+++}$ |
| | | (0.10;0.25)* | (0.09;0.31)* | | | |
| 5-comp. Weibull | 14 | 0.33 | 0.27 | 22,764 | 4 | $4^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32) | | | |
| 5-comp. Gamma | 14 | 0.36 | 0.29 | 22,758 | 5 | $5^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.31)* | | | |
| 4-comp. Gamma | 11 | 0.40 | 0.30 | 22,724 | 6 | $7^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.31)* | | | |
| 2-comp. Lognormal | 5 | 0.29 | 0.26 | 22,695 | 7 | $6^{+++}$ |
| | | (0.10;0.25)** | (0.09;0.31) | | | |
| 4-comp. Weibull | 11 | 0.40 | 0.29 | 22,691 | 8 | $8^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.31)* | | | |
| 4-comp. Exponential | 7 | 0.60 | 0.34 | 22,544 | 9 | $9^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.33)** | | | |
| 5-comp. Exponential | 9 | 0.59 | 0.36 | 22,541 | 10 | $10^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)** | | | |
| 3-comp. Burr | 11 | 0.30 | 0.38 | 22,477 | 11 | $11^{+++}$ |
| | | (0.09;0.26)*** | (0.08;0.32)** | | | |
| 3-comp. Weibull | 8 | 0.40 | 0.66 | 22,247 | 12 | $12^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.31)*** | | | |
| 5-comp. Fréchet | 14 | 0.67 | 0.56 | 22,240 | 13 | $13^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | | | |
| 3-comp. Gamma | 8 | 0.56 | 0.88 | 22,132 | 14 | $14^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.32)*** | | | |
| 4-comp. Fréchet | 11 | 0.68 | 0.66 | 22,056 | 15 | $16^{+++}$ |
| | | (0.09;0.26)*** | (0.08;0.32)*** | | | |
| 3-comp. Exponential | 5 | 0.64 | 1.00 | 22,025 | 16 | $15^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.33)*** | | | |
| 3-comp. Fréchet | 8 | 0.67 | 0.65 | 21,911 | 17 | $17^{+++}$ |
| | | (0.09;0.25)*** | (0.09;0.33)*** | | | |
| 2-comp. Burr | 7 | 0.80 | 0.76 | 21,614 | 22 | $22^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.30)*** | | | |
| Inv. Pareto-Burr-Pareto | 5 | 0.80 | 0.76 | 21,614 | 21 | $21^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.30)*** | | | |
| Inv. Pareto-Burr | 4 | 0.80 | 0.76 | 21,614 | 19 | $19^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | | | |
| Burr | 3 | 0.80 | 0.76 | 21,614 | 18 | $18^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.32)*** | | | |
| Burr-Pareto | 4 | 0.80 | 0.76 | 21,614 | 20 | $20^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.31)*** | | | |
| 5-comp. Burr | 19 | 0.80 | 0.76 | 21,614 | 23 | $24^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.31)*** | | | |
| Double-Pareto Lognormal | 4 | 1.04 | 1.36 | 21,592 | 24 | $23^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.33)*** | | | |
| Inv. Pareto-Lognormal-Pareto | 4 | 1.18 | 1.51 | 21,179 | 25 | $25^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.33)*** | | | |
| Inv. Pareto-Lognormal | 3 | 2.48 | 3.35 | 20,614 | 26 | $26^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.33)*** | | | |
| Right-Pareto Lognormal | 3 | 2.02 | 3.25 | 20,585 | 27 | $27^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.33)*** | | | |
| Lognormal-Pareto | 3 | 1.85 | 3.01 | 20,494 | 28 | $28^{+++}$ |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | (0.10;0.25)*** | (0.09;0.32)*** | | | |
| Left-Pareto Lognormal | 3 | 2.49 | 3.70 | 20,423 | 29 | $29^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.33)*** | | | |
| Lognormal | 2 | 2.35 | 3.68 | 20,407 | 30 | $30^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.33)*** | | | |
| Inv. Pareto-Fréchet-Pareto | 4 | 1.46 | 2.06 | 20,193 | 31 | $31^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.32)*** | | | |
| Inv. Pareto-Weibull-Pareto | 4 | 1.95 | 2.55 | 19,520 | 33 | $33^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.32)*** | | | |
| Weibull-Pareto | 3 | 1.95 | 2.55 | 19,520 | 32 | $32^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.33)*** | | | |
| Inv. Pareto-Gamma-Pareto | 4 | 2.06 | 2.75 | 19,441 | 35 | $35^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.32)*** | | | |
| Gamma-Pareto | 3 | 2.06 | 2.75 | 19,441 | 34 | $34^{+++}$ |
| | | (0.09;0.25)*** | (0.09;0.34)*** | | | |
| 2-comp. Weibull | 5 | 1.31 | 2.24 | 19,404 | 36 | $38^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | | | |
| Inv. Pareto-Exponential-Pareto | 3 | 2.20 | 3.02 | 19,394 | 38 | $37^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.31)*** | | | |
| Exponential-Pareto | 2 | 2.20 | 3.02 | 19,394 | 37 | $36^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.30)*** | | | |
| 2-comp. Fréchet | 5 | 1.49 | 2.25 | 19,272 | 39 | $39^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.32)*** | | | |
| 2-comp. Gamma | 5 | 2.30 | 3.50 | 16,675 | 40 | $40^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | | | |
| 2-comp. Exponential | 3 | 3.73 | 5.90 | 13,268 | 41 | $41^{+++}$ |
| | | (0.10;0.26)*** | (0.08;0.33)*** | | | |
| Fréchet | 2 | 7.68 | 12.96 | -6,681 | 42 | $42^{+++}$ |
| | | (0.09;0.26)*** | (0.08;0.33)*** | | | |
| Fréchet-Pareto | 3 | 7.68 | 12.96 | -6,681 | 44 | $43.5^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.32)*** | | | |
| Inv. Pareto-Fréchet | 3 | 7.68 | 12.96 | -6,681 | 44 | $43.5^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.33)*** | | | |
| Inv. Pareto-Weibull | 3 | 8.40 | 14.21 | -8,737 | 46 | $46^{+++}$ |
| | | (0.10;0.26)*** | (0.09;0.33)*** | | | |
| Weibull | 2 | 8.40 | 14.21 | -8,738 | 45 | $45^{+++}$ |
| | | (0.10;0.24)*** | (0.09;0.31)*** | | | |
| Inv. Pareto-Gamma | 3 | 15.94 | 24.26 | -43,526 | 47 | $47^{+++}$ |
| | | (0.09;0.26)*** | (0.08;0.34)*** | | | |
| Gamma | 2 | 15.94 | 24.27 | -43,533 | 48 | $48^{+++}$ |
| | | (0.09;0.24)*** | (0.08;0.30)*** | | | |
| Exponential | 1 | 32.58 | 56.49 | -139,654 | 49 | $49^{+++}$ |
| | | (0.10;0.25)*** | (0.08;0.32)*** | | | |
| Inv. Pareto-Exponential | 2 | 32.58 | 56.49 | -139,654 | 50 | $50^{+++}$ |
| | | (0.10;0.25)*** | (0.09;0.33)*** | | | |
| Pareto | 2 | 37.07 | 55.02 | -214,535 | 51 | $51^{+++}$ |
| | | (0.09;0.25)*** | (0.08;0.32)*** | | | |

**Notes:** All distributions fitted using Maximum Likelihood.
Values between parentheses report the 5th and 95th quantile of the parametric bootstrapped test statistic with 999 replications. ***, **, * indicate significance of this test at 1%, 5% and 10% respectively.
$^{+++}$, $^{++}$, $^{+}$ indicates the difference between this distribution's BIC and the first-ranked distribution in terms of BIC ($\Delta BIC$) providing strong evidence in favour of the first-ranked distribution ($\Delta BIC > 10$), moderate evidence ($6 < \Delta BIC \leq 10$) and weak evidence ($2 < \Delta BIC \leq 6$) respectively.
$_a$ Values multiplied by 100 for expositional purpose, $_b$ Values divided by 1,000 for expositional purpose.

Table 10: Distribution fits to the U.S. Census 2000 city size distribution.

| | | Goodness of fit | | Information Criteria | | |
|---|---|---|---|---|---|---|
| Distribution | Parms. | $T_a^0$ | $S_b^0$ | Loglike | $R_{AIC}$ | $R_{BIC}$ |
| 5-comp. Burr | 19 | 0.22<br>(0.33;0.87) | 0.02<br>(0.02;0.09) | -6,004 | 3 | $9^{+++}$ |
| 3-comp. Burr | 11 | 0.25<br>(0.33;0.85) | 0.02<br>(0.02;0.10) | -6,006 | 1 | $2^{++}$ |
| 4-comp. Burr | 15 | 0.32<br>(0.33;0.83) | 0.02<br>(0.02;0.09) | -6,008 | 2 | $5^{+++}$ |
| 5-comp. Lognormal | 14 | 0.58<br>(0.32;0.87) | 0.05<br>(0.02;0.10) | -6,016 | 6 | $6^{+++}$ |
| 4-comp. Lognormal | 11 | 0.60<br>(0.32;0.82) | 0.05<br>(0.02;0.09) | -6,016 | 5 | $4^{+++}$ |
| 3-comp. Lognormal | 8 | 0.62<br>(0.32;0.86) | 0.05<br>(0.02;0.09) | -6,017 | 4 | 1 |
| 5-comp. Gamma | 14 | 0.29<br>(0.32;0.84) | 0.03<br>(0.02;0.09) | -6,033 | 7 | $10^{+++}$ |
| 5-comp. Weibull | 14 | 0.38<br>(0.33;0.88) | 0.04<br>(0.03;0.10) | -6,037 | 9 | $12^{+++}$ |
| 2-comp. Lognormal | 5 | 0.71<br>(0.33;0.82) | 0.05<br>(0.02;0.09) | -6,044 | 8 | $3^{+++}$ |
| 2-comp. Burr | 7 | 0.87<br>(0.32;0.85)** | 0.09<br>(0.02;0.09)* | -6,056 | 10 | $7^{+++}$ |
| Right-Pareto Lognormal | 3 | 1.33<br>(0.31;0.84)*** | 0.17<br>(0.02;0.09)*** | -6,085 | 11 | $8^{+++}$ |
| Double-Pareto Lognormal | 4 | 1.39<br>(0.32;0.85)*** | 0.17<br>(0.02;0.09)*** | -6,085 | 12 | $11^{+++}$ |
| Inv. Pareto-Lognormal-Pareto | 4 | 1.76<br>(0.33;0.87)*** | 0.25<br>(0.02;0.09)*** | -6,135 | 14 | $14^{+++}$ |
| Lognormal-Pareto | 3 | 1.75<br>(0.33;0.86)*** | 0.25<br>(0.02;0.10)*** | -6,135 | 13 | $13^{+++}$ |
| 4-comp. Weibull | 11 | 0.69<br>(0.32;0.82) | 0.08<br>(0.02;0.09)* | -6,144 | 16 | $18^{+++}$ |
| Inv. Pareto-Lognormal | 3 | 1.90<br>(0.33;0.84)*** | 0.27<br>(0.02;0.09)*** | -6,152 | 17 | $16^{+++}$ |
| Lognormal | 2 | 1.89<br>(0.32;0.85)*** | 0.27<br>(0.02;0.09)*** | -6,152 | 15 | $15^{+++}$ |
| Left-Pareto Lognormal | 3 | 3.12<br>(0.98;1.93)*** | 0.42<br>(0.14;0.29)*** | -6,152 | 18 | $17^{+++}$ |
| 4-comp. Gamma | 11 | 0.99<br>(0.33;0.84)** | 0.11<br>(0.02;0.09)** | -6,163 | 19 | $19^{+++}$ |
| 5-comp. Fréchet | 14 | 1.73<br>(0.33;0.85)*** | 0.15<br>(0.02;0.09)*** | -6,172 | 21 | $21^{+++}$ |
| 4-comp. Fréchet | 11 | 1.57<br>(0.33;0.85)*** | 0.14<br>(0.02;0.09)*** | -6,174 | 20 | $20^{+++}$ |
| 5-comp. Exponential | 9 | 1.94<br>(0.32;0.86)*** | 0.11<br>(0.02;0.10)** | -6,260 | 22 | $23^{+++}$ |
| 3-comp. Fréchet | 8 | 1.61<br>(0.32;0.83)*** | 0.17<br>(0.02;0.09)*** | -6,261 | 23 | $22^{+++}$ |
| 4-comp. Exponential | 7 | 1.79<br>(0.32;0.84)*** | 0.14<br>(0.02;0.09)*** | -6,298 | 24 | $24^{+++}$ |
| Inv. Pareto-Burr | 4 | 2.17<br>(0.33;0.85)*** | 0.30<br>(0.03;0.09)*** | -6,370 | 26 | $26^{+++}$ |
| Inv. Pareto-Burr-Pareto | 5 | 2.17<br>(0.32;0.85)*** | 0.30<br>(0.02;0.10)*** | -6,370 | 28 | $28^{+++}$ |
| Burr-Pareto | 4 | 2.17<br>(0.32;0.85)*** | 0.30<br>(0.02;0.09)*** | -6,370 | 27 | $27^{+++}$ |
| Burr | 3 | 2.17<br>(0.32;0.84)*** | 0.30<br>(0.02;0.09)*** | -6,370 | 25 | $25^{+++}$ |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3-comp. Weibull | 8 | 1.71<br>(0.32;0.84)*** | 0.18<br>(0.02;0.10)*** | -6,393 | 29 | $29^{+++}$ |
| Inv. Pareto-Fréchet-Pareto | 4 | 3.05<br>(0.33;0.83)*** | 0.40<br>(0.02;0.09)*** | -6,530 | 30 | $30^{+++}$ |
| 3-comp. Gamma | 8 | 2.37<br>(0.33;0.85)*** | 0.25<br>(0.02;0.09)*** | -6,532 | 31 | $32^{+++}$ |
| 2-comp. Fréchet | 5 | 2.55<br>(0.32;0.84)*** | 0.32<br>(0.02;0.09)*** | -6,538 | 32 | $31^{+++}$ |
| 3-comp. Exponential | 5 | 2.76<br>(0.32;0.85)*** | 0.28<br>(0.02;0.09)*** | -6,633 | 33 | $33^{+++}$ |
| Inv. Pareto-Weibull-Pareto | 4 | 3.60<br>(0.32;0.86)*** | 0.48<br>(0.02;0.09)*** | -6,829 | 35 | $35^{+++}$ |
| Weibull-Pareto | 3 | 3.60<br>(0.32;0.88)*** | 0.48<br>(0.02;0.10)*** | -6,829 | 34 | $34^{+++}$ |
| Inv. Pareto-Gamma-Pareto | 4 | 3.87<br>(0.31;0.87)*** | 0.52<br>(0.02;0.09)*** | -6,848 | 37 | $39^{+++}$ |
| Gamma-Pareto | 3 | 3.87<br>(0.32;0.84)*** | 0.52<br>(0.02;0.10)*** | -6,848 | 36 | $37^{+++}$ |
| Inv. Pareto-Exponential-Pareto | 3 | 3.96<br>(0.32;0.82)*** | 0.54<br>(0.02;0.09)*** | -6,851 | 39 | $38^{+++}$ |
| Exponential-Pareto | 2 | 3.96<br>(0.32;0.85)*** | 0.54<br>(0.03;0.09)*** | -6,851 | 38 | $36^{+++}$ |
| 2-comp. Weibull | 5 | 2.96<br>(0.32;0.87)*** | 0.32<br>(0.03;0.09)*** | -6,920 | 40 | $40^{+++}$ |
| Fréchet-Pareto | 3 | 4.60<br>(0.32;0.85)*** | 0.64<br>(0.02;0.09)*** | -7,404 | 42 | $42.5^{+++}$ |
| Inv. Pareto-Fréchet | 3 | 4.60<br>(0.33;0.85)*** | 0.64<br>(0.02;0.09)*** | -7,404 | 42 | $42.5^{+++}$ |
| Fréchet | 2 | 4.60<br>(0.32;0.84)*** | 0.64<br>(0.02;0.09)*** | -7,404 | 41 | $41^{+++}$ |
| 2-comp. Gamma | 5 | 4.23<br>(0.32;0.86)*** | 0.54<br>(0.02;0.10)*** | -7,694 | 44 | $44^{+++}$ |
| 2-comp. Exponential | 3 | 7.16<br>(0.33;0.84)*** | 0.84<br>(0.03;0.09)*** | -8,488 | 45 | $45^{+++}$ |
| Inv. Pareto-Weibull | 3 | 8.31<br>(0.32;0.85)*** | 1.13<br>(0.02;0.09)*** | -9,030 | 47 | $47^{+++}$ |
| Weibull | 2 | 8.31<br>(0.32;0.82)*** | 1.13<br>(0.02;0.09)*** | -9,030 | 46 | $46^{+++}$ |
| Inv. Pareto-Gamma | 3 | 16.40<br>(0.33;0.84)*** | 2.26<br>(0.02;0.09)*** | -13,169 | 48 | $49^{+++}$ |
| Gamma | 2 | 16.42<br>(0.32;0.87)*** | 2.26<br>(0.02;0.10)*** | -13,171 | 49 | $48^{+++}$ |
| Exponential | 1 | 37.71<br>(0.32;0.87)*** | 5.58<br>(0.02;0.09)*** | -25,359 | 50 | $50^{+++}$ |
| Inv. Pareto-Exponential | 2 | 37.71<br>(0.32;0.85)*** | 5.58<br>(0.02;0.09)*** | -25,359 | 51 | $51^{+++}$ |
| Pareto | 2 | 41.69<br>(0.33;0.83)*** | 5.06<br>(0.02;0.09)*** | -31,612 | 52 | $52^{+++}$ |

**Notes:** All distributions fitted using Maximum Likelihood.
Values between parentheses report the 5th and 95th quantile of the parametric bootstrapped test statistic with 999 replications. ***, **, * indicate significance of this test at 1%, 5% and 10% respectively.
+++, ++, + indicates the difference between this distribution's BIC and the first-ranked distribution in terms of BIC ($\Delta BIC$) providing strong evidence in favour of the first-ranked distribution ($\Delta BIC > 10$), moderate evidence ($6 < \Delta BIC \leq 10$) and weak evidence ($2 < \Delta BIC \leq 6$) respectively.
$_a$ Values multiplied by 100 for expositional purpose, $_b$ Values divided by 1,000 for expositional purpose.

Table 11: Decomposition of procentual welfare gains from a reduction in variable trade costs $\tau^{ij} = 3 \rightarrow (\tau^{ij})' = 1$.

| Distribution | Parms. | $ln\frac{U_i'}{U_i}$ | $ln\frac{\tau_{ij}'}{\tau_{ij}}$ | $ln\frac{M_i'}{M_i}$ | $ln\frac{1-G(\omega_{ij}^*)'}{1-G(\omega_{ij}^*)}$ | $ln\frac{\tilde{\omega}(\omega_{ij}^*)'}{\tilde{\omega}(\omega_{ij}^*)}$ | $ln\frac{\lambda_{ij}'}{\lambda_{ij}}$ |
|---|---|---|---|---|---|---|---|
| Pareto | 2 | - | 1.10 | - | - | - | - |
| | | (-0.00;0.00)*** | (1.10;1.10) | (-0.22;-0.22)*** | (-0.00;0.00)*** | (0.00;0.00)*** | (-0.88;-0.88)*** |
| Weibull | 2 | 0.15 | 1.10 | -0.16 | 0.12 | 1.35 | -2.26 |
| | | (0.15;0.15)*** | (1.10;1.10) | (-0.16;-0.16)*** | (0.12;0.13)*** | (1.30;1.41)*** | (-2.32;-2.21)*** |
| Inv. Pareto-Weibull | 3 | 0.15 | 1.10 | -0.16 | 0.12 | 1.35 | -2.26 |
| | | (0.15;0.15)*** | (1.10;1.10) | (-0.16;-0.16)*** | (0.12;0.13)*** | (1.30;1.41)*** | (-2.32;-2.21)*** |
| Left-Pareto Lognormal | 3 | 0.16 | 1.10 | -0.17 | 0.15 | 0.60 | -1.51 |
| | | (0.16;0.17)*** | (1.10;1.10) | (-0.17;-0.17)*** | (0.15;0.15)*** | (0.58;0.62)*** | (-1.53;-1.49)*** |
| Inv. Pareto-Lognormal | 3 | 0.17 | 1.10 | -0.17 | 0.15 | 0.58 | -1.49 |
| | | (0.16;0.17)*** | (1.10;1.10) | (-0.17;-0.17)*** | (0.15;0.15)*** | (0.56;0.60)*** | (-1.51;-1.47)*** |
| Lognormal | 2 | 0.17 | 1.10 | -0.17 | 0.15 | 0.53 | -1.44 |
| | | (0.17;0.17)*** | (1.10;1.10) | (-0.17;-0.17)*** | (0.15;0.15)*** | (0.51;0.55)*** | (-1.46;-1.42)*** |
| Right-Pareto Lognormal | 3 | 0.18 | 1.10 | -0.18 | 0.17 | 0.28 | -1.19 |
| | | (0.18;0.19)** | (1.10;1.10) | (-0.19;-0.18) | (0.17;0.18) | (0.23;0.33)** | (-1.24;-1.13)** |
| 2-comp. Weibull | 5 | 0.18 | 1.10 | -0.13 | 0.11 | 0.57 | -1.46 |
| | | (0.18;0.18)*** | (1.10;1.10) | (-0.14;-0.13)*** | (0.11;0.11)*** | (0.56;0.59)*** | (-1.48;-1.45)*** |
| 3-comp. Weibull | 8 | 0.19 | 1.10 | -0.19 | 0.18 | 0.25 | -1.16 |
| | | (0.18;0.19)*** | (1.10;1.10) | (-0.19;-0.19)*** | (0.18;0.19)*** | (0.25;0.26)*** | (-1.17;-1.15)*** |
| 4-comp. Weibull | 11 | 0.19 | 1.10 | -0.18 | 0.17 | 0.22 | -1.12 |
| | | (0.19;0.19)*** | (1.10;1.10) | (-0.18;-0.17)*** | (0.16;0.17)*** | (0.21;0.23)*** | (-1.13;-1.11)*** |
| 5-comp. Weibull | 14 | 0.19 | 1.10 | -0.18 | 0.18 | 0.22 | -1.12 |
| | | (0.19;0.19)** | (1.10;1.10) | (-0.19;-0.18) | (0.17;0.18) | (0.21;0.23)*** | (-1.14;-1.11)*** |
| Empirical | 0 | 0.19 | 1.10 | -0.18 | 0.18 | 0.20 | -1.10 |
| 4-comp. Lognormal | 11 | 0.19 | 1.10 | -0.18 | 0.18 | 0.20 | -1.10 |
| | | (0.19;0.19) | (1.10;1.10) | (-0.19;-0.18) | (0.17;0.18) | (0.18;0.22) | (-1.13;-1.08) |
| 5-comp. Lognormal | 14 | 0.19 | 1.10 | -0.19 | 0.18 | 0.20 | -1.10 |
| | | (0.19;0.19) | (1.10;1.10) | (-0.19;-0.18) | (0.17;0.19) | (0.17;0.22) | (-1.12;-1.07) |
| 2-comp. Lognormal | 5 | 0.19 | 1.10 | -0.17 | 0.17 | 0.23 | -1.13 |
| | | (0.19;0.19) | (1.10;1.10) | (-0.18;-0.17)*** | (0.16;0.17)*** | (0.22;0.25)*** | (-1.15;-1.12)*** |
| 3-comp. Lognormal | 8 | 0.19 | 1.10 | -0.18 | 0.18 | 0.19 | -1.09 |
| | | (0.19;0.19) | (1.10;1.10) | (-0.19;-0.18) | (0.17;0.18) | (0.16;0.22) | (-1.12;-1.06) |
| Lognormal-Pareto | 3 | 0.22 | 1.10 | -0.22 | 0.22 | 0.02 | -0.90 |
| | | (0.20;0.21)*** | (1.10;1.10) | (-0.22;-0.20)*** | (0.20;0.22)*** | (0.04;0.14)*** | (-1.04;-0.93)*** |
| Burr | 3 | - | 1.10 | - | - | - | - |
| | | (0.20;0.21)*** | (1.10;1.10) | (-0.21;-0.19)*** | (0.19;0.21)*** | (0.03;0.12)*** | (-1.02;-0.92)*** |
| 2-comp. Burr | 7 | - | 1.10 | - | - | - | - |
| | | (0.19;0.20) | (1.10;1.10) | (-0.20;-0.18) | (0.17;0.20) | (0.10;0.22) | (-1.12;-1.00) |
| 3-comp. Burr | 11 | - | 1.10 | - | - | - | - |
| | | (0.19;0.20)** | (1.10;1.10) | (-0.21;-0.18) | (0.18;0.21) | (0.08;0.20) | (-1.11;-0.97) |
| 4-comp. Burr | 15 | - | 1.10 | - | - | - | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | (0.19;0.21)** | (1.10;1.10) | (-0.21;-0.18)** | (0.18;0.21)** | (0.07;0.20)** | (-1.10;-0.96)** |
| 5-comp. Burr | 19 | - | 1.10 | - | - | - | - |
| | | (0.19;0.21)*** | (1.10;1.10) | (-0.22;-0.19)*** | (0.18;0.22)*** | (0.05;0.19)*** | (-1.09;-0.94)*** |
| Burr-Pareto | 4 | - | 1.10 | - | - | - | - |
| | | (0.20;0.21)*** | (1.10;1.10) | (-0.21;-0.19)*** | (0.19;0.21)*** | (0.02;0.12)*** | (-1.02;-0.91)*** |
| Double-Pareto Lognormal | 4 | - | 1.10 | - | - | - | - |
| | | (0.20;0.22)*** | (1.10;1.10) | (-0.20;-0.19)*** | (0.19;0.20)*** | (0.02;0.09)*** | (-0.98;-0.90)*** |
| Fréchet | 2 | - | 1.10 | - | - | - | - |
| | | (0.22;0.22)*** | (1.10;1.10) | (-0.14;-0.08)*** | (0.08;0.14)*** | (0.00;0.01)*** | (-0.89;-0.88)*** |
| 2-comp. Fréchet | 5 | - | 1.10 | - | - | - | - |
| | | (0.22;0.22)*** | (1.10;1.10) | (-0.15;-0.10)*** | (0.10;0.15)*** | (0.00;0.01)*** | (-0.89;-0.88)*** |
| 3-comp. Fréchet | 8 | - | 1.10 | - | - | - | - |
| | | (0.22;0.22)*** | (1.10;1.10) | (-0.15;-0.11)** | (0.11;0.15)** | (0.00;0.01)*** | (-0.89;-0.88)*** |
| 4-comp. Fréchet | 11 | - | 1.10 | - | - | - | - |
| | | (0.22;0.22)*** | (1.10;1.10) | (-0.15;-0.11)** | (0.11;0.15)** | (0.00;0.01)*** | (-0.89;-0.88)*** |
| 5-comp. Fréchet | 14 | - | 1.10 | - | - | - | - |
| | | (0.22;0.22)*** | (1.10;1.10) | (-0.15;-0.10)*** | (0.10;0.15)** | (0.00;0.01)*** | (-0.89;-0.88)*** |
| Fréchet-Pareto | 3 | - | 1.10 | - | - | - | - |
| | | (0.22;0.22)*** | (1.10;1.10) | (-0.14;-0.08)*** | (0.08;0.14)*** | (0.00;0.01)*** | (-0.89;-0.88)*** |
| Inv. Pareto-Burr | 4 | - | 1.10 | - | - | - | - |
| | | (0.20;0.22)*** | (1.10;1.10) | (-0.21;-0.19)*** | (0.19;0.21)*** | (0.02;0.11)*** | (-1.00;-0.90)*** |
| Inv. Pareto-Burr-Pareto | 5 | - | 1.10 | - | - | - | - |
| | | (0.20;0.22)*** | (1.10;1.10) | (-0.21;-0.19)*** | (0.19;0.20)*** | (0.02;0.11)*** | (-1.00;-0.90)*** |
| Inv. Pareto-Fréchet | 3 | - | 1.10 | - | - | - | - |
| | | (0.22;0.22)*** | (1.10;1.10) | (-0.14;-0.08)*** | (0.08;0.14)** | (0.00;0.01)*** | (-0.89;-0.88)*** |
| Inv. Pareto-Fréchet-Pareto | 4 | - | 1.10 | - | - | - | - |
| | | (0.21;0.22)*** | (1.10;1.10) | (-0.19;-0.18) | (0.18;0.19)** | (0.01;0.07)*** | (-0.96;-0.89)*** |
| Inv. Pareto-Lognormal-Pareto | 4 | - | 1.10 | - | - | - | - |
| | | (0.21;0.22)*** | (1.10;1.10) | (-0.20;-0.18) | (0.18;0.20)*** | (0.01;0.08)*** | (-0.97;-0.89)*** |
| Inv. Pareto-Weibull-Pareto | 4 | - | 1.10 | - | - | - | - |
| | | (0.21;0.22)*** | (1.10;1.10) | (-0.18;-0.16)** | (0.16;0.18) | (0.00;0.05)*** | (-0.93;-0.88)*** |
| Weibull-Pareto | 3 | - | 1.10 | - | - | - | - |
| | | (0.21;0.22)*** | (1.10;1.10) | (-0.18;-0.16)** | (0.16;0.18) | (0.00;0.05)*** | (-0.93;-0.88)*** |

**Notes:** Values between parentheses report the 5th and 95th quantile of the parametric bootstrapped statistics with 999 replications. ***, **, * indicate the rejection of a signifcant overlap of the parametric bootstrapped statistic with the empirical statistic at 1%, 5% and 10% respectively.

# Appendix B Motivation and identification of generative processes for mixture models

FMMs can be utilized in two ways. First, they can be used as a semi-parametric, flexible approximation of the overall distribution, which is the case in this paper. Second, they are model-based clustering methods when a certain distribution is imposed (Fop et al., 2018; Grün, 2018). While both applications rely on the idea that discrete subpopulations define the overall distribution, the semi-parametric approximation does not claim to identify these subpopulations. In this appendix, we conceptualize possible Data Generative Processes (DGPs) for FMMs based on theoretical and empirical work in the economics literature. We then elaborate on the identification difficulties/opportunities of the underlying mixture components in the context of productivity distributions.

## B.1 Generative processes

Many economic models rely on the assumption that the firm size distribution originates from firm dynamics in productivity (see for instance Hopenhayn (1992); Luttmer (2007); Rossi-Hansberg and Wright (2007); Costantini and Melitz (2008); Arkolakis (2016)). In this section, we will use a simplified version of such productivity dynamics for explanatory purposes. Consider productivity dynamics specified as a first-order autoregressive process:

$$ln\omega_{bt} = c + \rho ln\omega_{bt-1} + \eta_{bt}, \tag{1}$$

where $\eta_{bt}$ is a white noise process with zero mean and constant variance $\sigma^2$.

There exists empirical evidence arguing that productivity dynamics, and therefore the resulting productivity distributions, are endogenous to exporting (De Loecker, 2013), importing (Kasahara and Rodrigue, 2008), innovation (Aw et al., 2011), management practices (Bloom and Reenen, 2011; Caliendo et al., 2020), …Overall, there are "many sources of heterogeneity that support the idea of discrete subpopulations likely to differ in important characteristics …" (Perline (2005),p.80). In the case of exporting, the endogenous evolution of productivity results in an exporting productivity premium. This can empirically be observed from the standard textbook comparison of cross-sectional productivity densities between exporting and non-exporting firms (see Figure 5). Building on equation 1, a simplified version of the empirical specification to identify such exporting productivity premium, and replicate Figure 5, is essentially a specifically parametrized FMM:

$$\begin{aligned} ln\omega_{bt} &= \alpha_0 + \beta_0 EXP_b + \alpha_1 ln\omega_{bt-1} + \beta_1 EXP_b \times ln\omega_{bt-1} + \eta_{bt} \\ &= EXP_b \left[\beta_0 + \beta_1 ln\omega_{bt-1}\right] + (1 - EXP_b)\left[\alpha_0 + \alpha_1 ln\omega_{bt-1}\right] + \eta_{bt}, \end{aligned} \tag{2}$$

with $EXP_b$ a dummy variable that takes the value 1 when the firm $b$ is an exporter and 0 otherwise.

Whereas the components are identified by means of an exporter dummy variable in this example, FMMs are a semi-parametric specification that remain agnostic about the (possibly multiple) determinants of the unobserved components and allow the data to determine these components:[1]

$$ln\omega_{bt} = \sum_{i=1}^{I} \mathbb{I}_b^i \left[\beta_0^i + \beta_1^i ln\omega_{bt-1}\right] + \eta_{bt}. \tag{3}$$

[1]Note that, for simplicity, we specify the variance to be constant between components. FMMs in the main analysis allow for the variance to differ between components.

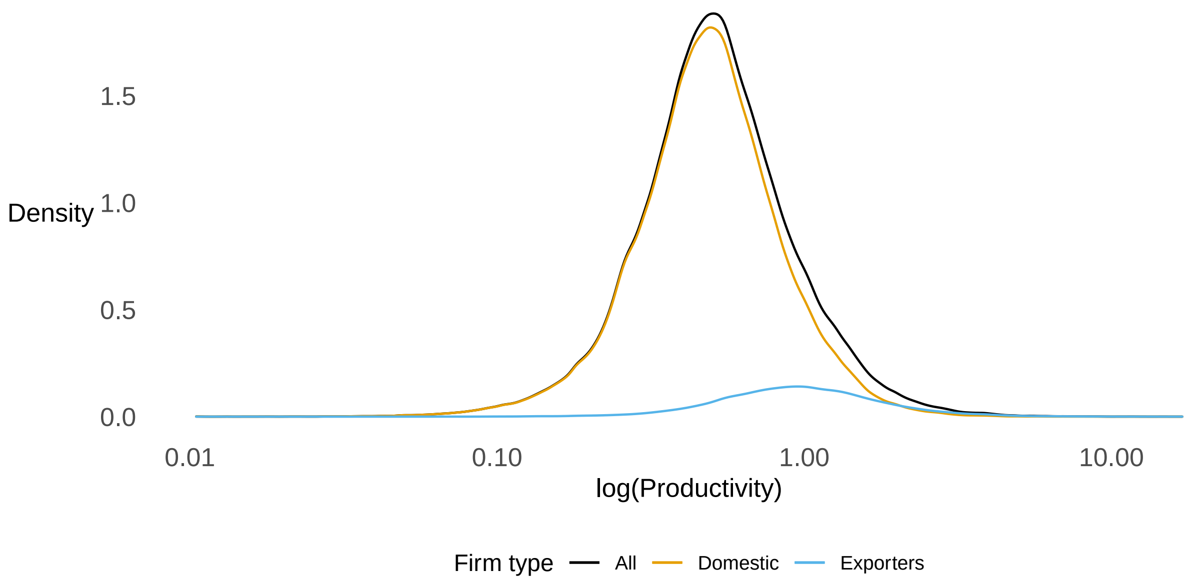


Figure 5: Productivity density of Portuguese firm productivity in 2006 for all, exporting- and non-exporting firms.

**Notes:** Productivity is measured as domestic sales (relative to the mean) to the power of $1/(\sigma - 1)$ with $\sigma$, the elasticity of substitution between varieties, set to four.

### B.2 Identification

As stated before, the use of FMM's can focus on the semi-parametric, flexible approximation of the overall distribution or on model-based clustering. This paper purely focuses on the semi-parametric approximation, with good reason. First, we take no à-priori stance on distributional specification.[2] Second, even if one is willing to assume distributional specification such as the Lognormal, the underlying components remain unidentifiable in the current setting. As the overall distribution is unimodal (see Figure 5), there is a large overlap between the underlying individual densities. These individual densities will therefore be poorly identified. Indeed, Figure 6 displays the posterior probability distribution for each component of the fitted 4-component Lognormal mixture from the main text. Whereas well-identified components have a large weight near zero and 1, average probabilities lie close to 0.25 in our case, and are therefore not well-identified. While the overall distribution can be closely approximated, the large overlap of individual densities results in a large uncertainty on which observation can be assigned to which density. Neither the parameter estimates used to characterize the clusters nor the partitions derived can therefore be uniquely determined, rendering the interpretation of results in terms of clustering futile (Follmann and Lambert, 1991; Hennig, 2000; Grün, 2018; Grün and Leisch, 2008).

[2]The empirical evidence in this paper seems to favor a Lognormal specification. This can be motivated from two perspectives. From the perspective of overall fit, a mixture of (log-) normal distributions with sufficient components is assumed to be able to approach all distributions (McLachlan and Peel, 2000). From a generative perspective for individual components, the Lognormal distribution is the realization of applying the Central Limit Theorem (CLT) in the log domain: firm heterogeneity will approximately be Lognormal if it is the multiplicative product of many independent random variables. Whereas firm heterogeneity reduces to firm-level productivity in the Melitz (2003)-model, it has been argued to be multi-dimensional when taking into consideration for instance the product dimension (Bernard et al., 2009) or uncertainty in demand and/or supply (see (De Loecker, 2011; Bas et al., 2017; Sager and Timoshenko, 2019; Gandhi et al., 0) . . . )

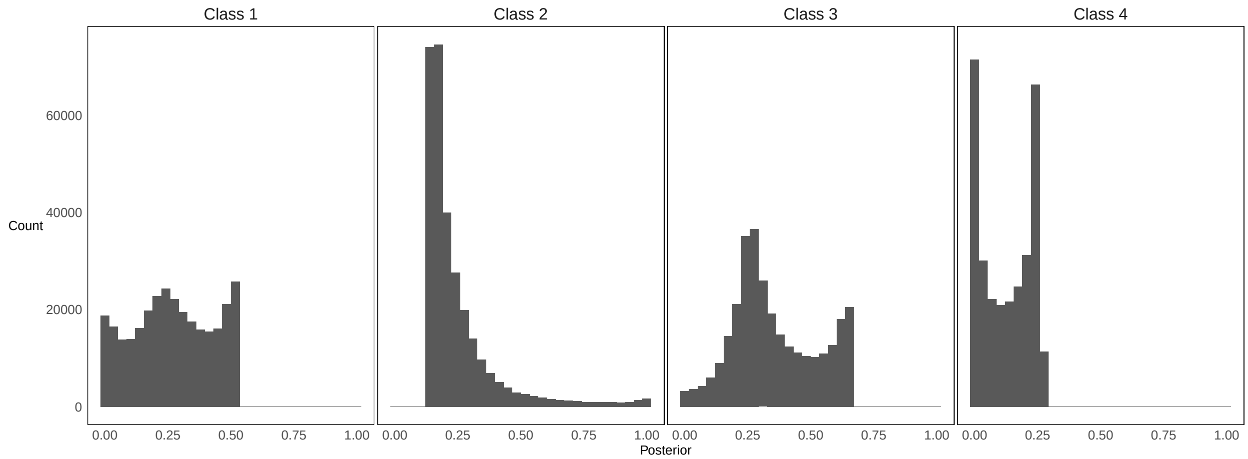


Figure 6: Posterior probability distribution for each component of the 4-components Lognormal mixture.

Future research might resolve the identifiability problem relying on panel rather than cross-sectional data. The problem as specified now is a problem in levels (the cross-section), where it appears there is insufficient distance between different components for them to be identified. From empirical evidence, however, it can be deduced that the different components likely originate from differences in growth rates (Kasahara and Rodrigue, 2008; Aw et al., 2011; De Loecker, 2013; Caliendo et al., 2020). Tracking the growth rates of individual firms over time might allow for the variation needed to identify the components of the overall distribution.

This observation can be easily illustrated using simulated data. Building on the example of the previous paragraph, imagine $ln\omega_{bt}$ follows an AR(1)-process with an exporting productivity premium of 20%:

$$ln\omega_{bt} = 1 + 1.2 \times EXP_b + 0.7 \times ln\omega_{bt-1} + \eta_{bt},$$

with $\eta_{bt} \sim \mathcal{N}(0, 0.3)$. We simulate this evolution for 200 exporters ($EXP_b = 1$) and 800 purely domestic businesses over 10 years.[3] The firm densities of the simulated data will look similar to Figure 5, with two densities largely overlapping but the exporter productivity density located on the right of domestic firms density.

If we fit, as in our main analysis, a FMM on the cross-sectional data of a selected (the first) year, we obtain a familiar posterior probability distribution (see Figure 7). Individual clusters are not well-identified. Exploiting the panel dimension of the data,[4] however, results in well-identified

[3] When simulating, we allow for a run-in period of 90 years.

[4] Specifically, the EM estimation procedure is adapted to take into account panel data. The component probabilities in our main analysis are specified over the complete data (eq. 9):

$$\pi_{bi}^{(s)} = E\left[z_{bi}|\omega_b, \mathbf{\Psi}^{(s-1)}\right] = \frac{\pi_i^{(s-1)} m_i(\omega_b|\boldsymbol{\theta}_i^{(s-1)})}{\sum_{i=1}^{I} \pi_i^{(s-1)} m_i(\omega_b|\boldsymbol{\theta}_i^{(s-1)})}.$$

When working with panel data, we adapt this specification to take into account the time dimension:

$$\pi_{bi}^{(s)} = E\left[z_{bi}|\omega_{bt}, \mathbf{\Psi}^{(s-1)}\right] = \frac{\pi_i^{(s-1)} \prod_{t=1}^{T} m_{it}(\omega_{bt}|\boldsymbol{\theta}_i^{(s-1)})}{\sum_{i=1}^{I} \pi_i^{(s-1)} \prod_{t=1}^{T} m_{it}(\omega_{bt}|\boldsymbol{\theta}_i^{(s-1)})}.$$

Note that the probabilities are specified to be constant over time, meaning that we do not allow for regime switching components. As can be observed in Figure 8, the posterior probabilities predominantly take the values zero or one. Once components are well-identified, one can try to determine which mechanisms motivate the existence of FMMs from a generative perspective.

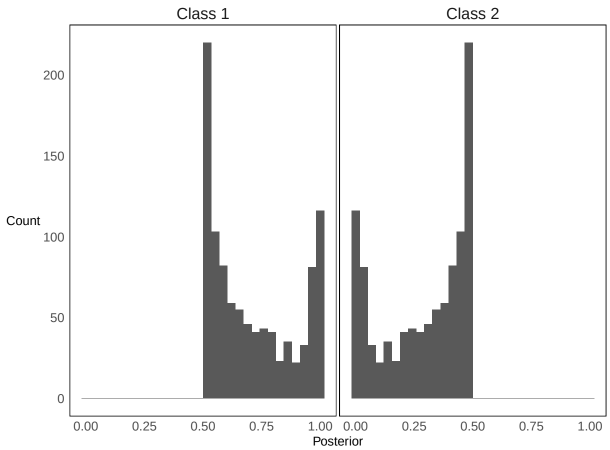


Figure 7: *Cross-sectional* posterior probability distribution for each component of the simulated 2-components normal mixture.

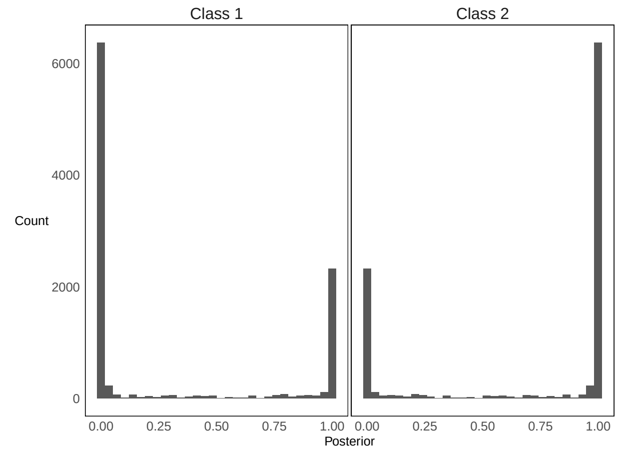


Figure 8: *Panel* posterior probability distribution for each component of the simulated 2-components normal mixture.

in this exercise.

# Appendix C Heterogeneous firms model

This appendix provides a detailed description of the heterogeneous firms models relied upon in the paper. We follow (Dewitte, 2020) in presenting a firm heterogeneous open economy model of Melitz (2003) with a finite number of firms. The model features Constant Elasticity of Substitution (CES)-demand and monopolistic competition between a finite number of firms who ignore their aggregate impact (Dixit and Stiglitz, 1977; Krugman, 1980; di Giovanni and Levchenko, 2012), while remaining agnostic on the parametric specification of firm-level heterogeneity. For the number of firms going to infinity, the model is equivalent to the Melitz (2003)- model.

## C.1 Setup

**Demand** Consumer preferences in country $j \in J$ are defined over a finite number of horizontally differentiated varieties ($\varpi \in \Omega^i$) originating from country $i \in I$ and are assumed to take the Constant Elasticity of Substitution (CES) utility ($U$) form

$$U^j = \left( \sum_{i=1}^{I} \sum_{\varpi \in \Omega^i} q^{ij} \left( \varpi \right)^{\frac{\sigma-1}{\sigma}} d\varpi \right)^{\frac{\sigma}{\sigma-1}}, \tag{4}$$

with $\sigma$ the elasticity of substitution between varieties. Utility maximization defines the optimal consumption and expenditure decisions over the individual varieties

$$\frac{q^{ij}(\varpi)}{Q^j} = \left[ \frac{p^{ij}(\varpi)}{P^j} \right]^{-\sigma}, \tag{5}$$

where the set of varieties consumed is considered as an aggregate good $Q \equiv U$ and $P$ is the CES aggregate price index.

**Supply** There is a finite number of businesses ($b \in B$) which choose to supply a distinct horizontally-differentiated variety. They are heterogeneous in terms of their productivity $\omega_b \in [0, \infty]$ drawn from the unconditional Cumulative Distribution Function (CDF) $G(\omega_b)$ after paying a fixed cost $f^{ie}$ in terms of production factor $L^i$ to enter the market.[5] There is zero probability of firm death.[6] Supply of the production factor to the individual firm is perfectly elastic, so that firms are effectively price ($W^i$) takers on the input markets. Once active, firms from country $i$ have to pay a fixed cost $f^{ij}$ to produce goods destined for country j. The cost function of the firm involves a fixed production cost, iceberg trade costs $\tau^{ij} > 1$ and a constant marginal costs that depends on its productivity: $f^{ij} + \left( \frac{\tau^{ij} q^{ij}}{\omega} \right) W^i$. Profit maximization of the firm, then:

$$\begin{aligned} \max_{q^{ij}} \pi^{ij} &= \max_{q^{ij}} \left[ p^{ij} q^{ij} - \left( f^{ij} - \frac{\tau^{ij} q^{ij}}{\omega} \right) W^i \right] \\ &= \max_{q^{ij}} \left[ \left( q^{ij} \right)^{\frac{\sigma-1}{\sigma}} \left( Q^j \right)^{\frac{1}{\sigma}} P^j - \left( f^{ij} - \frac{\tau^{ij} q^{ij}}{\omega} \right) W^i \right], \end{aligned} \tag{6}$$

[5] As $\omega_b$ is the sole heterogeneity component identifying individual firms, we drop the subscript $b$ in further derivations.

[6] The static specification in which there is zero probability of firm death follows most of the international trade literature.

results in an optimal quantity produced:

$$\frac{\partial \pi^{ij}}{\partial q^{ij}} = 0$$
$$\Leftrightarrow$$
$$\frac{\sigma-1}{\sigma}\left(q^{ij}\right)^{-\frac{1}{\sigma}}\left(Q^{j}\right)^{\frac{1}{\sigma}}P^{j} = \frac{\tau^{ij}W^{i}}{\omega}$$
$$\Leftrightarrow$$
$$q^{ij} = \left(\frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^{i}}{\omega}\right)^{-\sigma} Q^{j}\left(P^{j}\right)^{\sigma}. \tag{7}$$

and an equilibrium price as a constant markup over marginal costs $p^{ij} = \frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^{i}}{\omega}$:

$$\left(\frac{q^{ij}}{(Q^{j})}\right)^{\frac{-1}{\sigma}} P^{j} = p^{ij}$$
$$p^{ij} = \frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^{i}}{\omega}. \tag{8}$$

The realized revenue expression for firms from country $i$ selling in destination $j$ at time t can then be expressed as:

$$\begin{aligned} x^{ij} = p^{ij}q^{ij} &= \left(q^{ij}\right)^{\frac{\sigma-1}{\sigma}}\left(Q^{j}\right)^{\frac{1}{\sigma}}P^{j} \\ &= \left(\frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^{i}}{\omega}\right)^{1-\sigma} Q^{j}\left(P^{j}\right)^{\sigma} \end{aligned} \tag{9}$$

## C.2 Operating decisions

In line with (Dixit and Stiglitz, 1977; Krugman, 1980; di Giovanni and Levchenko, 2012), we assume that the marginal firm ignores the impact its own production level on the aggregate economy. The necessary productivity levels for serving each market are then determined by the zero cutoff profit conditions.

$$
\begin{aligned}
\pi^{ij} = 0 &= p^{ij}q^{ij} - \left(f^{ij} - \frac{\tau^{ij}q^{ij}}{\omega^{ij*}}\right)W^i, \\
&= \left(\frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^i}{\omega^{ij*}}\right)^{1-\sigma} Q^j \left(P^j\right)^{\sigma} - f^{ij}W^i - \left(\frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^i}{\omega^{ij*}}\right)^{-\sigma} Q^j \left(P^j\right)^{\sigma} \frac{\tau^{ij}}{\omega^{ij*}}W^i, \\
&= \left(\frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^i}{\omega^{ij*}}\right)^{1-\sigma} Q^j \left(P^j\right)^{\sigma} - f^{ij}W^i - \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \left(\frac{\tau^{ij}W^i}{\omega^{ij*}}\right)^{1-\sigma} Q^j \left(P^j\right)^{\sigma}, \\
&= \left(1 - \frac{\sigma-1}{\sigma}\right)\left(\frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^i}{\omega^{ij*}}\right)^{1-\sigma} Q^j \left(P^j\right)^{\sigma} - f^{ij}W^i, \\
&\Leftrightarrow \\
\sigma f^{ij}W^i &= \left(\frac{\sigma}{\sigma-1}\frac{\tau^{ij}W^i}{\omega^{ij*}}\right)^{1-\sigma} Q^j \left(P^j\right)^{\sigma}. \qquad (10)
\end{aligned}
$$

Combining the zero cutoff profit conditions allows us to write the export cutoff as a function of a foreign domestic productivity cutoff, variable and fixed costs and the wages:

$$\omega^{ij*} = \left(\frac{W^i}{W^j}\right)^{\frac{\sigma}{\sigma-1}} \left(\frac{f^{ij}}{f^{jj}}\right)^{\frac{1}{\sigma-1}} \left(\frac{\tau^{ij}}{\tau^{jj}}\right)\omega^{jj*}. \qquad (11)$$

Similarly, we can combine the zero cutoff profit conditions from a single origin country, linking the domestic and export productivity cutoffs:

$$\omega^{ij*} = \frac{\tau^{ij}}{\tau^{ii}}\left(\frac{P^j}{P^i}\right)^{\frac{\sigma}{1-\sigma}} \left(\frac{Q^i}{Q^j}\frac{f^{ij}}{f^{ii}}\right)^{\frac{1}{\sigma-1}} \omega^{ii*}. \qquad (12)$$

In this paper, we focus on parameter values such that there is, in line with empirical evidence, selection into exporting ($\omega^{ij*} > \omega^{ii*}$). This implies

- A large fixed cost of exporting relative to the fixed cost of production. The revenue required to cover the fixed export cost is then large relative to the revenue required to cover the fixed production cost, implying that only firms of high productivity find it profitable to serve both markets.

- A high home price index relative to the foreign price index, and a large home market relative to the foreign market. Only high productivity firms receive enough revenue in the relatively small and competitive foreign market to cover the fixed cost of exporting.

- Variable trade costs increase the exporting productivity cutoff relative to the zero-profit productivity cutoff by increasing prices and reducing revenue in the export market.

The equilibrium value of these cutoffs are uniquely determined by the free entry condition, requiring the probability of successful entry times the expected future value of entry conditional upon successful entry to equal the sunk entry cost:

$$\sum_{j=1}^{J} \mathbb{E}\left[\pi^{ij} | \omega > \omega^{ij*}\right] = f^{ie} W^i$$

$$\sum_{j=1}^{J} \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\left(\omega > \omega^{ij*}\right) \pi^{ij} = f^{ie} W^i$$

$$\sum_{j=1}^{J} \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\left(\omega > \omega^{ij*}\right) \left[\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{\tau^{ij} W^i}{\omega}\right)^{1-\sigma} Q^j \left(P^j\right)^{\sigma} - f^{ij} W^i\right] = f^{ie} W^i$$

$$\sum_{j=1}^{J} f^{ij} W^i \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\left(\omega > \omega^{ij*}\right) \left[\left(\frac{\omega}{\omega^{ij*}}\right)^{\sigma-1} - 1\right] = f^{ie} W^i$$

$$\sum_{j=1}^{J} f^{ij} \left[\left(\omega^{ij*}\right)^{1-\sigma} \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\left(\omega > \omega^{ij*}\right) \omega^{\sigma-1} - \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\left(\omega > \omega^{ij*}\right) \omega^{0}\right] = f^{ie}$$

$$\sum_{j=1}^{J} f^{ij} \left[\left(\omega^{ij*}\right)^{1-\sigma} m^{\sigma-1}_{\omega^{ij*}} - m^{0}_{\omega^{ij*}}\right] = f^{ie}, \tag{13}$$

where we denote by $m_y^r$ the y-bounded, r-th sample moment of the productivity distribution. For the number of firms going to infinity, the law of large numbers kicks in such that we replace these sample moments with their continuous equivalent $\left(\mu^r(y) = \int_y^\infty \omega^r g(\omega) d\omega\right)$, providing us with the well-known continuous free-entry equation as specified by (Melitz, 2003).

Using the relation between productivity cutoffs (eq. 11), the free entry condition (eq. 13) determines a unique equilibrium values of these cutoffs.[7] Thus, a parametrization of the Melitz (2003)-model in relation to firm heterogeneity relies solely on the bounded (by the respective productivity cutoffs) 0th and $(\sigma - 1)$th moments of the productivity distribution (Nigai, 2017; Dewitte, 2020).

### C.3 Aggregation

Summing equation 9 across all active firms, we obtain an expression for aggregate trade between country $i$ and $j$:

$$X^{ij} = \left(\frac{\sigma}{\sigma-1}\tau^{ij} W^i\right)^{1-\sigma} Q^j \left(P^j\right)^{\sigma} M^{ie} m^{\sigma-1}_{\omega^{ij*}} \tag{14}$$

The number of successful entrants $\left[1 - G(\omega^{ii*})\right] M^{ie}$ is specified as the ratio of aggregate over average revenue:

$$M^i = \left[1 - G(\omega^{ii*})\right] M^{ie} = \frac{X^i}{\mathbb{E}\left[x^i\right]}. \tag{15}$$

[7] Sufficient conditions for this equilibrium to exist are that the term in brackets of equation (13) is (i) finite and (ii) a decreasing function of the cutoffs (Melitz, 2003, p.1704). The second condition corresponds to $\frac{g(x)x}{1-G(x)}$ increasing to infinity on $(0, \infty)$.

We can rewrite this number of firms, using the free entry condition, goods and labor market clearing $\left(X^i = W^i L^i\right)$, as a function of exogenous variables:

$$M^i = \frac{W^i L^i}{\sigma\left(\frac{f^{ie}}{1-G(\omega^{ii*})} + \sum_{j=1}^{J} \frac{1-G(\omega^{ij*})}{1-G(\omega^{ii*})} f^{ij}\right) W^i}$$
$$= \frac{L^i}{\sigma\left(\frac{f^{ie}}{1-G(\omega^{ii*})} + \sum_{j=1}^{J} \frac{1-G(\omega^{ij*})}{1-G(\omega^{ii*})} f^{ij}\right)}. \tag{16}$$

Assuming a two-country symmetric economy and setting the wage of the composite factor as the numeraire, welfare can be calculated as the inverse of the price index

$$\mathbb{W}^i = (P^i)^{-1}. \tag{17}$$

The price index can be deduced from equation 14:

$$P^j = \left[\left(\frac{\sigma}{\sigma-1}\tau^{ij} W^i\right)^{1-\sigma} \frac{1}{\lambda_{ij}} \frac{M^i}{1-G(\omega^{ii*})} m_{\omega^{ij*}}^{\sigma-1}\right]^{\frac{1}{1-\sigma}}, \tag{18}$$

where we denote the bilateral trade share by $\lambda^{ij} = \frac{X^{ij}}{X^j}$.

The percentage changes in welfare from a change in variable trade costs ($\boldsymbol{\tau} \rightarrow \boldsymbol{\tau}'$) can then written as:

$$\begin{aligned} 100 \times ln\frac{(\mathbb{W}^i)'}{\mathbb{W}^i} &= 100 \times -ln\frac{(P^i)'}{P^i} \\ &= 100 \times -ln\frac{(P^j)'}{P^j} \\ &= 100 \times -\left[ln\frac{(\tau^{ij})'}{(\tau^{ij})} - \frac{1}{\sigma-1}\left(ln\frac{(M^i)'}{M^i} - ln\frac{1-G(\omega^{ii*})'}{1-G(\omega^{ii*})} + ln\frac{(m_{\omega^{ij*}}^{\sigma-1})'}{m_{\omega^{ij*}}^{\sigma-1}} - ln\frac{(\lambda^{ij})'}{\lambda^{ij}}\right)\right] \\ &= 100 \times -\left[ln\frac{(\tau^{ij})'}{(\tau^{ij})} - \frac{1}{\sigma-1}\left(ln\frac{(M^i)'}{M^i} - ln\frac{(m_{\omega^{ij*}}^{0})'}{m_{\omega^{ij*}}^{0}} + ln\frac{(m_{\omega^{ij*}}^{\sigma-1})'}{m_{\omega^{ij*}}^{\sigma-1}} - ln\frac{(\lambda^{ij})'}{\lambda^{ij}}\right)\right]. \end{aligned} \tag{19}$$

## C.4 Parametrization

In order to parametrize the previously described model, we need to parametrize two statistics related to the productivity distribution: the 0th and $(\sigma-1)$the y-bounded moments of the *productivity* distribution (Nigai, 2017). As described in (Dewitte, 2020), this corresponds to the 0th and 1st y-bounded moments of the *sales* distribution if the parametric distribution is stable under power-law transformations.

Assuming a parametric distribution and under the assumption of an *infinite* number of firms, we can calculate the necessary analytical expressions using the distributional parameters from our empirical analysis to capture heterogeneity. This is the standard approach in the literature. Following (Nigai, 2017; Dewitte, 2020), we can also capture heterogeneity directly from the empirical, *finite*, data. To allow comparison between GFT obtained assuming a parametric distribution and GFT obtained from the finite data, we perform a parametric bootstrap. This parametric bootstrap generates a range of finite sample estimates under the hypothesis that the observed data is generated by a certain parametric distribution (Dewitte, 2020).

### C.4.1 Continuum of firms

When there are an infinite number of firms, the parametrization of the heterogeneity distribution consists of calculating the y-bounded 0th and 1st population moments of the sales distribution:

$$\mu_y^r = \int_y^\infty x^r g(x) dx. \tag{20}$$

The analytical expressions of these parametric implied population moments are gathered in Table 4 and 5 for all distributions considered. As bounded moments are not generally available, the mathematical elaboration on how to obtain these expressions can be found in the section D.

### C.4.2 Finite number of firms

Under the assumption of a finite number of firms in the economy, the parametrization of the model consists of calculating the y-bounded 0th and 1st moment of the sales distribution:

$$m_y^r = \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\left(x > y\right) x^r. \tag{21}$$

These moments can easily be retrieved if the data is available. To allow comparison between GFT obtained assuming a parametric distribution and GFT obtained from the finite data, we perform a parametric bootstrap. This parametric bootstrap generates a range of finite sample estimates under the hypothesis that the observed data is generated by a certain parametric distribution:

1. Assume B i.i.d. random variables with distribution $G(\cdot|\boldsymbol{\theta})$, with empirical finite sample moments $m_y^r$ for $r = 0, 1$, as specified in equation 21 and corresponding $GFT_B$;
2. Estimate the parameters $\boldsymbol{\theta}$ of the distribution using MLE, calculate the parametric plug-in population moments as specified in equation 20, $\hat{\mu}^r(y|\hat{\boldsymbol{\theta}})$ for $r = 0, 1$, and corresponding $G\hat{F}T(\hat{\boldsymbol{\theta}})$;
3. $H_0 : GFT = G\hat{F}T(\hat{\boldsymbol{\theta}})$;
4. Draw N bootstrap samples of size B from $G(\cdot|\hat{\boldsymbol{\theta}})$;
5. For each sample of the parametric distribution, calculate the bootstrapped sample moments $(m_y^r)^*$ and calculate the corresponding $GFT_B^*$.[8]

[8]Note that we do not re-fit the parametric distribution to the bootstrap sample. The vastness of the dataset at our availability in the empirical section results both in a large computational burden but also a very precise estimation of the distribution parameters. The influence of not refitting the parametric distribution to the bootstrap sample is therefore negligent.

6. The p-value for the left-, and right-tailed test is then respectively specified as:

$$\hat{p}_l = \frac{1}{N+1}\left[\sum_{n=1}^{N} \mathbb{I}\left(GFT_B^* \geq GFT_B\right) + 1\right]; \qquad \hat{p}_r = \frac{1}{N+1}\left[\sum_{n=1}^{N} \mathbb{I}\left(GFT_B^* \leq GFT_B\right) + 1\right]. \tag{22}$$

The bootstrap exercise should therefore be interpreted as 'the likelihood of observing GFT as small or as large as $GFT_B$ under the null hypothesis that the observed data originates from the parametric distribution $G(\cdot|\boldsymbol{\theta})$', allowing us to evaluate whether the distributional assumption provides a good fit to calculate GFT within the proposed model.

When calculating the bounded sample moments, complications can arise related to the lower bound $y$. This lower bound is ex ante unknown, can take values not observed in the data and/or resides in an unrepresentative part of the finite dataset.[9] We address each issue below and argue that these complications have little influence on our results.

1. $y$ can take values within the boundaries of the data but are not observed. We use the 'approx-fun' interpolation function of the R base distribution to approximate the statistics for such lower bounds.[10] As the calculation of Gains From Trade (GFT) relies on domestic cutoffs residing in the dense part of the productivity distribution, the influence of interpolation is negligible.

2. $y$ can take values below the lowest observed value in the data ($y < x_{min}$):

$$\mu_y^r = \underbrace{\sum \mathbb{I}\left(y < x < x_{min}\right) x^r}_{\text{unobserved}} + \underbrace{\frac{1}{B}\sum_{b=1}^{B} \mathbb{I}\left(x \geq x_{min}\right) x^r}_{observed}. \tag{23}$$

   The error arising from neglecting the unobserved part of the distribution is likely small as (i) the smallest observation $x_{min}$ in our dataset is rather small, (ii) the density in the unobserved part is most likely very low and (iii) the relative weight of the observations in the unobserved part is small (see also Figure 1).

3. As the presented model is a stylized model, it is conceivable firms produce below the model's implied zero-profit productivity cutoff, for instance when there is a positive expectation of future profits (Impullitti et al., 2013). This can explain very low observed productivity values, but will result in an unrepresentative left tail of the distribution (the lower the actual zero-profit productivity cutoff, the more firms will have a positive expectation of future profits and the denser the left tail of the distribution will be). This issue affects both the nonparametric and parametric estimates, as the parametric distribution is fitted to the observed distribution. Also in this case, however, provided the low density in the left tail of the distribution and the low relative weight of the observations in the left tail, the influence of this issue is likely small.

[9]We thank Gonzague Vannoorenberghe for pointing this out.
[10]All code available on request.

# Appendix D Analytical expressions of $\mu_y^r$

## D.1 Pareto

$$
\begin{aligned}
\mu_y^r &= \int_y^\infty x^r \frac{kx_{min}^k}{x^{k+1}} dx \\
&= kx_{min}^k \frac{-y^{r-k}}{r-k} \qquad \text{if } k > r
\end{aligned} \tag{24}
$$

## D.2 Inverse Pareto

$$
\begin{aligned}
\mu_y^r &= \int_y^{x_{max}} x^r \frac{kx_{max}^{-k}}{x^{-k+1}} dx \\
&= kx_{max}^{-k} \frac{x_{max}^{r+k} - y^{r+k}}{r+k}
\end{aligned} \tag{25}
$$

## D.3 Lognormal

$$
\begin{aligned}
\mu_y^r &= \int_y^\infty x^r \frac{1}{xVar\sqrt{2\pi}} e^{-(lnx-\mu)^2/2Var^2} dx \\
&= \int_y^\infty e^{rlnx} \frac{1}{xVar\sqrt{2\pi}} e^{-(lnx-\mu)^2/2Var^2} dx
\end{aligned} \tag{26}
$$

Note that

$$
\begin{aligned}
rlnx - (lnx-\mu)^2/2Var^2 &= \frac{2Var^2 rlnx - (lnx)^2 - \mu^2 + 2\mu lnx}{2Var^2} \\
&= -\frac{(lnx)^2 - 2(Var^2 r + \mu)lnx + \left((Var^2 r + \mu)\right)^2 - (Var^2 r + \mu)^2 + \mu^2}{2Var^2} \\
&= -\frac{\left[lnx - \left(Var^2 r + \mu\right)\right]^2}{2Var^2} + \frac{(Var^2 r + \mu)^2 - \mu^2}{2Var^2} \\
&= -\frac{\left[lnx - \left(Var^2 r + \mu\right)\right]^2}{2Var^2} + \frac{r\left(rVar^2 + 2\mu\right)}{2}
\end{aligned}
$$

so that

$$
\begin{aligned}
\mu_y^r &= e^{\frac{r\left(rVar^2+2\mu\right)}{2}} \int_y^\infty \frac{1}{xVar\sqrt{2\pi}} e^{-\frac{\left[lnx-\left(Var^2 r+\mu\right)\right]^2}{2Var^2}} dx \\
&\text{let } z = \frac{lnx - \left(rVar^2 + \mu\right)}{Var},\ dz = \frac{dx}{xVar} \\
&= e^{\frac{r\left(rVar^2+2\mu\right)}{2}} \int_{\frac{lny-\left(rVar^2+\mu\right)}{Var}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dx \\
&= e^{\frac{r\left(rVar^2+2\mu\right)}{2}} \left[1 - \Phi\left(\frac{lny - \left(rVar^2 + \mu\right)}{Var}\right)\right]
\end{aligned} \tag{27}
$$

### D.4 Weibull[11]

$$
\begin{aligned}
\mu_y^r &= \int_y^{\infty} x^r \frac{k}{s} \left(\frac{x}{s}\right)^{k-1} e^{-\left(\frac{x}{s}\right)^k} dx \\
&\qquad \text{let } z = \left(\frac{x}{s}\right)^k, dz = \frac{k}{s}\left(\frac{x}{s}\right)^{k-1} dx \\
&\qquad \text{s.t. } x = s z^{\frac{1}{k}} \\
&= \int_{\left(\frac{y}{s}\right)^k}^{\infty} s^r z^{\frac{r}{k}} e^{-z} dz \\
&= s^r \int_{\left(\frac{y}{s}\right)^k}^{\infty} z^{\left(\frac{r}{k}+1\right)-1} e^{-z} dz \\
&= s^r \Gamma\left(\frac{r}{k}+1, \left(\frac{y}{s}\right)^k\right) \qquad (28)
\end{aligned}
$$

where $\Gamma(,)$ denotes the upper incomplete gamma function.

### D.5 Fréchet

$$
\begin{aligned}
\mu_y^r &= \int_y^{\infty} x^r \frac{k}{s} \left(\frac{x}{s}\right)^{-1-k} e^{-\left(\frac{x}{s}\right)^{-k}} dx \\
&\qquad \text{let } z = \left(\frac{x}{s}\right)^{-k}, dz = \frac{-k}{s}\left(\frac{x}{s}\right)^{-k-1} dx \\
&\qquad \text{s.t. } x = s z^{-\frac{1}{k}} \\
&= -\int_{\left(\frac{y}{s}\right)^{-k}}^{0} s^r z^{-\frac{r}{k}} e^{-z} dz, \qquad \text{if } k > 0 \\
&= \int_0^{\left(\frac{y}{s}\right)^{-k}} s^r z^{-\frac{r}{k}} e^{-z} dz \\
&= s^r \int_0^{\left(\frac{y}{s}\right)^{-k}} z^{1-\left(\frac{r}{k}\right)-1} e^{-z} dz \\
&= s^r \left[1 - \Gamma\left(1 - \frac{r}{k}, \left(\frac{y}{s}\right)^{-k}\right)\right] \qquad \text{if } k > r \qquad (29)
\end{aligned}
$$

[11]The bounded moments of the exponential distribution are obtained setting k=1.

### D.6 Burr

$$\mu_y^r = \int_y^\infty x^r \frac{\frac{kc}{s}\left(\frac{x}{s}\right)^{c-1}}{\left(1+\left(\frac{x}{s}\right)^c\right)^{k+1}} dx$$

$$\text{let } z = \left(\frac{x}{s}\right)^c, dz = \frac{c}{s}\left(\frac{x}{s}\right)^{c-1} dx$$

$$\text{s.t. } x = sz^{\frac{1}{c}}$$

$$= \int_{\left(\frac{y}{s}\right)^c}^\infty s^r z^{\frac{r}{c}} \frac{k}{(1+z)^{k+1}} dz, \qquad \text{if } c > 0$$

$$= s^r k \int_{\left(\frac{y}{s}\right)^c}^\infty z^{\frac{r}{c}} \frac{1}{(1+z)^{k+1}} dz$$

$$= s^r k \int_{\left(\frac{y}{s}\right)^c}^\infty z^{\left(\frac{r}{c}+1\right)-1} \frac{1}{(1+z)^{k+1}} dz$$

$$= s^r k \int_{\left(\frac{y}{s}\right)^c}^\infty z^{\left(\frac{r}{c}+1\right)-1} \frac{1}{(1+z)^{k+1}} dz$$

$$= s^r k \left[\int_0^\infty z^{\left(\frac{r}{c}+1\right)-1} \frac{1}{(1+z)^{k+1}} dz - \int_0^{\left(\frac{y}{s}\right)^c} z^{\left(\frac{r}{c}+1\right)-1} \frac{1}{(1+z)^{k+1}} dz\right]$$

$$u = \frac{z}{1+z}, du = \frac{1}{(1+z)^2}$$

$$z = \frac{u}{1-u}$$

$$= s^r k \left[\int_0^1 \left(\frac{u}{1-u}\right)^{\left(\frac{r}{c}+1\right)-1} \frac{1}{\left(1+\frac{u}{1-u}\right)^{k-1}} du - \int_0^{\frac{\left(\frac{y}{s}\right)^c}{1+\left(\frac{y}{s}\right)^c}} \left(\frac{u}{1-u}\right)^{\left(\frac{r}{c}+1\right)-1} \frac{1}{\left(1+\frac{u}{1-u}\right)^{k-1}} du\right]$$

$$= s^r k \left[\int_0^1 u^{\left(\frac{r}{c}+1\right)-1} (1-u)^{k-1-\left(\frac{r}{c}+1\right)+1} du\right.$$

$$\left. - \int_0^{\frac{\left(\frac{y}{s}\right)^c}{1+\left(\frac{y}{s}\right)^c}} \int_0^{\left(\frac{y}{s}\right)^c} u^{\left(\frac{r}{c}+1\right)-1} (1-u)^{k-1-\left(\frac{r}{c}+1\right)+1} du\right]$$

$$= s^r k \left[\int_0^1 u^{\left(\frac{r}{c}+1\right)-1} (1-u)^{k-\left(\frac{r}{c}+1\right)} du - \int_0^{\frac{\left(\frac{y}{s}\right)^c}{1+\left(\frac{y}{s}\right)^c}} u^{\left(\frac{r}{c}+1\right)-1} (1-u)^{k-\left(\frac{r}{c}+1\right)} duz\right]$$

$$= s^r k \left[\boldsymbol{B}\left(\frac{r}{c}+1, k-\frac{r}{c}\right) - \boldsymbol{B}\left(\frac{\left(\frac{y}{s}\right)^c}{1+\left(\frac{y}{s}\right)^c}; \frac{r}{c}+1, k-\frac{r}{c}\right)\right] \qquad \text{if } c > r, kc > r \tag{30}$$

where $\boldsymbol{B}(a,b)$ stands for the beta function, while $\boldsymbol{B}(x,a,b)$ stands for the lower incomplete beta function with upper bound $x$.

### D.7 Generalized Gamma[12]

$$\mu_y^r = \int_y^\infty x^r \frac{c}{s^k \Gamma(\frac{k}{c})} x^{k-1} e^{-(\frac{x}{s})^c} dx$$

$$\text{let } z = \left(\frac{x}{s}\right)^c, dz = \frac{c}{s}\left(\frac{x}{s}\right)^{c-1} dx$$

$$\text{s.t. } x = s z^{\frac{1}{c}}$$

$$= \int_{(\frac{y}{s})^c}^\infty s^r \frac{z^{\frac{r}{c}}}{\Gamma(\frac{k}{c})} \left(\frac{s z^{\frac{1}{c}}}{s}\right)^{(k-1)-(c-1)} e^{-z} dz, \qquad \text{if } c > 0$$

$$= \frac{s^r}{\Gamma(\frac{k}{c})} \int_{(\frac{y}{s})^c}^\infty z^{\frac{r+k}{c}-1} e^{-z} dz$$

$$= \frac{s^r}{\Gamma(\frac{k}{c})} \Gamma\left(\frac{r+k}{c}, \left(\frac{y}{s}\right)^c\right) \tag{31}$$

### D.8 Finite Mixture Model

The statistics for a Finite Mixture Model can easily be obtained from the calculated statistics for the underlying individual distributions on which the mixture consists. For a mixture of the form:

$$g(x|\boldsymbol{\Psi}) = \sum_{i=1}^{I} \pi_i m_i(x|\boldsymbol{\theta}_i), \qquad \pi_i \geq 0, \quad \sum_{i=1}^{I} \pi_i = 1, \tag{32}$$

we obtain, due to its additivity and applying the sum rule in integration:

$$\mu_y^r = \int_y^\infty x^r g(x|\boldsymbol{\Psi}) dx = \int_y^\infty x^r \sum_{i=1}^{I} \pi_i m_i(x|\boldsymbol{\theta_i}) dx = \sum_{i=1}^{I} \pi_i \int_y^\infty x^r m_i(x) dx = \sum_{i=1}^{I} \pi_i (\mu_i)_y^r. \tag{33}$$

### D.9 Piecewise composite

$$\mu_y^r = \int_y^\infty x^r g(x|\boldsymbol{\theta}) dx$$

$$= \begin{cases} \frac{\alpha_1}{1+\alpha_1+\alpha_2} \frac{(\mu_1)_y^r - (\mu_1)_{c_1}^r}{M_1(c_1)} + \frac{1}{1+\alpha_1+\alpha_2} \frac{(\mu_2)_{c_1}^r - (\mu_2)_{c_2}^r}{M_2(c_2) - M_2(c_1)} + \frac{\alpha_2}{1+\alpha_1+\alpha_2} \frac{(\mu_3)_y^r}{1-M_3(c_2)} & \text{if } \quad 0 < y \leq c_2 \\ \frac{1}{1+\alpha_1+\alpha_2} \frac{(\mu_2)_y^r - (\mu_2)_{c_2}^r}{M_2(c_2) - M_2(c_1)} + \frac{\alpha_2}{1+\alpha_1+\alpha_2} \frac{(\mu_3)_{c_2}^r}{1-M_3(c_2)} & \text{if } \quad c_1 < y \leq c_2 \\ \frac{\alpha_2}{1+\alpha_1+\alpha_2} \frac{(\mu_3)_y^r}{1-M_3(c_2)} & \text{if } \quad c_2 < y < \infty \end{cases} \tag{34}$$

[12]The bounded moments of the Gamma distribution are obtained setting c=1.

### D.10 Right-Pareto Lognormal

$$\begin{aligned}
\mu_y^r &= \int_y^\infty x^r k_2 x^{-k_2-1} e^{k_2\mu + \frac{k_2^2 Var^2}{2}} \Phi\left(\frac{lnx - \mu - k_2 Var^2}{Var}\right) dx \\
&= k_2 e^{k_2\mu + \frac{k_2^2 Var^2}{2}} \int_y^\infty x^{\sigma - k_2 - 2} \Phi\left(\frac{lnx - \mu - k_2 Var^2}{Var}\right) dx \\
&\qquad dv = x^{\sigma - k_2 - 2} dx, v = \frac{x^{\sigma - k_2 - 1}}{\sigma - k_2 - 1} \\
&\qquad u = \Phi\left(\frac{lnx - \mu - k_2 Var^2}{Var}\right), du = d\Phi\left(\frac{lnx - \mu - k_2 Var^2}{Var}\right) \\
&= k_2 e^{k_2\mu + \frac{k_2^2 Var^2}{2}} \left[\frac{x^{\sigma - k_2 - 1}}{\sigma - k_2 - 1} \Phi\left(\frac{lnx - \mu - k_2 Var^2}{Var}\right)\right]_y^\infty \\
&\quad - k_2 e^{k_2\mu + \frac{k_2^2 Var^2}{2}} \int_y^\infty \frac{x^{\sigma - k_2 - 1}}{\sigma - k_2 - 1} d\Phi\left(\frac{lnx - \mu - k_2 Var^2}{Var}\right) \\
&= k_2 e^{k_2\mu + \frac{k_2^2 Var^2}{2}} \left[0 - \frac{x_{ij^*}^{\sigma - k_2 - 1}}{\sigma - k_2 - 1} \Phi\left(\frac{lny - \mu - k_2 Var^2}{Var}\right)\right] \\
&\quad - k_2 e^{k_2\mu + \frac{k_2^2 Var^2}{2}} \int_y^\infty \frac{x^{\sigma - k_2 - 1}}{\sigma - k_2 - 1} \frac{1}{x Var \sqrt{2\pi}} e^{-\frac{\left[lnx - \mu - k_2 Var^2\right]^2}{2 Var^2}} dx
\end{aligned} \tag{35}$$

The last integral resembles the bounded moment condition of the Lognormal distribution solved earlier with moment $(r - k_2)$ and mean $(\mu + k_2 Var^2)$ so that

$$\begin{aligned}
\mu_y^r = &-k_2 e^{k_2\mu + \frac{k_2^2 Var^2}{2}} \frac{x_{ij^*}^{\sigma - k_2 - 1}}{\sigma - k_2 - 1} \Phi\left(\frac{lny - \mu - k_2 Var^2}{Var}\right) \\
&- \frac{k_2 e^{k_2\mu + \frac{k_2^2 Var^2}{2}}}{r - k_2} e^{\frac{(r-k_2)\left((r-k_2)Var^2 + 2(\mu + k_2 Var^2)\right)}{2}} \Bigg[1 \\
&\qquad -\Phi\left(\frac{lny - \left((r - k_2)Var^2 - (\mu + k_2 Var^2)\right)}{Var}\right)\Bigg]
\end{aligned} \tag{36}$$

Note that

$$\begin{aligned}
e^{k_2\mu + \frac{k_2^2 Var^2}{2} + \frac{(r-k_2)\left((r-k_2)Var^2 + 2(\mu + k_2 Var^2)\right)}{2}} \\
e^{\frac{2k_2\mu + k_2^2 Var^2 + (r-k_2)\left[rVar^2 + 2\mu + k_2 Var^2\right]}{2}} \\
e^{\frac{2k_2\mu + k_2^2 Var^2 + r^2 Var^2 + 2\mu r + k_2 r Var^2 - k_2 r Var^2 - 2\mu k_2 + k_2^2 Var^2}{2}} \\
e^{\frac{r^2 Var^2 + 2\mu r}{2}}
\end{aligned}$$

so that we get

$$
\begin{aligned}
\mu_y^r = &-k_2 e^{k_2\mu + \frac{k_2^2 Var^2}{2}} \frac{x_{ij^*}^{\sigma-k_2-1}}{\sigma-k_2-1} \Phi\left(\frac{lny-\mu-k_2Var^2}{Var}\right) \\
&- \frac{k_2}{r-k_2} e^{\frac{r^2Var^2+2\mu r}{2}} \Phi^c\left(\frac{lny-rVar^2-\mu}{Var}\right)
\end{aligned} \tag{37}
$$

### D.11 Left-Pareto Lognormal

$$
\begin{aligned}
\mu_y^r &= \int_y^\infty x^r x^{k_1-1} e^{-k_1\mu + \frac{k_1^2Var^2}{2}} \Phi^c\left(\frac{lnx-\mu+k_1Var^2}{Var}\right) dx \\
&= k_1 e^{k_2\mu + \frac{k_2^2Var^2}{2}} \left[\left(\frac{-(y)^{\sigma-k_2-1}}{\sigma-k_2-1}\right) - e^{-k_1\mu+\frac{k_1^2Var^2}{2}} \int_y^\infty x^{\sigma-2+k_1} \Phi\left(\frac{lnx-\mu+k_1Var^2}{Var}\right) dx\right] \\
&= -k_1 e^{-k_1\mu+\frac{k_1^2Var^2}{2}} \frac{x_{ij^*}^{\sigma+k_1-1}}{\sigma+k_1-1} \Phi^c\left(\frac{lny-\mu+k_1Var^2}{Var}\right) \\
&\quad + \frac{k_1}{r+k_1} e^{\frac{r^2Var^2+2\mu r}{2}} \Phi^c\left(\frac{lny-rVar^2+\mu}{Var}\right)
\end{aligned} \tag{38}
$$

### D.12 Double-Pareto Lognormal

$$
\begin{aligned}
\mu_y^r = &\frac{k_2k_1}{k_2+k_1} \int_y^\infty x^r x^{-k_2-1} e^{k_2\mu+\frac{k_2^2Var^2}{2}} \Phi\left(\frac{lnx-\mu-k_2Var^2}{Var}\right) dx \\
&+ \frac{k_2k_1}{k_2+k_1} \int_y^\infty x^r x^{k_1-1} e^{-k_1\mu+\frac{k_1^2Var^2}{2}} \Phi^c\left(\frac{lnx-\mu+k_1Var^2}{Var}\right) dx \\
&= -\frac{k_2k_1}{k_2+k_1} e^{k_2\mu+\frac{k_2^2Var^2}{2}} \frac{x_{ij^*}^{\sigma-k_2-1}}{\sigma-k_2-1} \Phi\left(\frac{lny-\mu-k_2Var^2}{Var}\right) \\
&\quad - \frac{k_2k_1}{k_2+k_1} \frac{1}{r-k_2} e^{\frac{r^2Var^2+2\mu r}{2}} \Phi^c\left(\frac{lny-rVar^2-\mu}{Var}\right) \\
&- \frac{k_2k_1}{k_2+k_1} e^{-k_1\mu+\frac{k_1^2Var^2}{2}} \frac{x_{ij^*}^{\sigma+k_1-1}}{\sigma+k_1-1} \Phi\left(\frac{lny-\mu+k_1Var^2}{Var}\right) \\
&\quad - \frac{k_2k_1}{k_2+k_1} \frac{1}{r+k_1} e^{\frac{r^2Var^2+2\mu r}{2}} \Phi^c\left(\frac{lny-rVar^2-\mu}{Var}\right)
\end{aligned} \tag{39}
$$

## Appendix References

Arkolakis, C. (2016). A Unified Theory of Firm Selection and Growth. *The Quarterly Journal of Economics 131*(1), 89.

Aw, B. Y., M. J. Roberts, and D. Y. Xu (2011, June). R&D Investment, Exporting, and Productivity Dynamics. *American Economic Review 101*(4), 1312–44.

Bas, M., T. Mayer, and M. Thoenig (2017). From micro to macro: Demand, supply, and heterogeneity in the trade elasticity. *Journal of International Economics 108*, 1 – 19.

Bernard, A. B., S. J. Redding, and P. K. Schott (2009). Products and productivity. *The Scandinavian Journal of Economics 111*(4), 681–709.

Bloom, N. and J. V. Reenen (2011). Chapter 19 - human resource management and productivity. Volume 4 of *Handbook of Labor Economics*, pp. 1697–1767. Elsevier.

Caliendo, L., G. Mion, L. D. Opromolla, and E. Rossi-Hansberg (2020). Productivity and organization in portuguese firms. *Journal of Political Economy forthcoming.*

Costantini, J. and M. Melitz (2008). The dynamics of firm-level adjustment to trade liberalization. *The organization of firms in a global economy 4*, 107–141.

De Loecker, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica 79*(5), 1407–1451.

De Loecker, J. (2013). Detecting learning by exporting. *American Economic Journal: Microeconomics 5*(3), 1–21.

Dewitte, R. (2020). From Heavy-tailed Micro to Macro: On the characterization of ffirm-level heterogeneity and its aggregation properties. mimeo, Ghent University.

di Giovanni, J. and A. A. Levchenko (2012). Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy 120*(6), 1083–1132.

Dixit, A. K. and J. E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *The American Economic Review 67*(3), 297–308.

Follmann, D. A. and D. Lambert (1991). Identifiability of finite mixtures of logistic regression models. *Journal of Statistical Planning and Inference 27*(3), 375–381.

Fop, M., T. B. Murphy, et al. (2018). Variable selection methods for model-based clustering. *Statistics Surveys 12*, 18–65.

Gandhi, A., S. Navarro, and D. A. Rivers (0). On the identification of gross output production functions. *Journal of Political Economy forthcoming.*

Grün, B. (2018). Model-based clustering. arXiv preprint 1807.01987, arXiv.

Grün, B. and F. Leisch (2008). Identifiability of finite mixtures of multinomial logit models with varying and fixed effects. *Journal of classification 25*(2), 225–247.

Hennig, C. (2000). Identifiablity of models for clusterwise linear regression. *Journal of classification 17*(2).

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica 60*(5), 1127–1150.

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.

Kasahara, H. and J. Rodrigue (2008). Does the use of imported intermediates increase productivity? plant-level evidence. *Journal of Development Economics 87*(1), 106–118.

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

Luttmer, E. G. (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics 122*(3), 1103–1144.

McLachlan, G. J. and D. Peel (2000). *Finite mixture models*. New York: Wiley Series in Probability and Statistics.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Nigai, S. (2017). A tale of two tails: Productivity distribution and the gains from trade. *Journal of International Economics 104*, 44–62.

Perline, R. (2005). Strong, weak and false inverse power laws. *Statistical Science 20*(1), 68–88.

Rossi-Hansberg, E. and M. L. J. Wright (2007, December). Establishment size dynamics in the aggregate economy. *American Economic Review 97*(5), 1639–1666.

Sager, E. and O. A. Timoshenko (2019). The double emg distribution and trade elasticities. *Canadian Journal of Economics/Revue canadienne d'économique 52*(4), 1523–1557.