Journal of Mathematical Economics 60 (2015) 1–8

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

# On the modeling of size distributions when technologies are complex[☆]

Jakub Growiec[*]

*Warsaw School of Economics, Department of Quantitative Economics, and Narodowy Bank Polski, Poland*

ARTICLE INFO

*Article history:*
Received 22 January 2015
Accepted 5 June 2015
Available online 22 June 2015

*Keywords:*
Technological productivity distribution
Stochastic R&D
CES
Weibull distribution
Lognormal distribution
Power law

ABSTRACT

The study considers a stochastic R&D process where the invented production technologies consist of a large number $n$ of complementary components. The degree of complementarity is captured by the elasticity of substitution of the CES aggregator function. Drawing from the Central Limit Theorem and the Extreme Value Theory we find, under very general assumptions, that the cross-sectional distributions of technological productivity are well-approximated either by the lognormal, Weibull, or a novel "CES/Normal" distribution, depending on the underlying elasticity of substitution between technology components. We find the tail of the "CES/Normal" distribution to be fatter than the Weibull tail but qualitatively thinner than the Pareto (power law) one. We also numerically assess the rate of convergence of the true technological productivity distribution to the theoretical limit with $n$ as fast in the body but slow in the tail.



## 1. Introduction

Most technologies used nowadays are complex in the sense that the production processes (and products themselves) consist of a large number of components which might interact with each other in complementary ways (e.g. Kremer, 1993; Blanchard and Kremer, 1997; Jones, 2011). Based on this insight, the current paper assumes that the total productivity of any given technology is functionally dependent on the individual productivities of its $n$ components as well as the elasticity of substitution between them, $\sigma$. This functional relationship is captured by the CES aggregator function. The stochastic R&D process which invents new complex technologies is in turn assumed to consist in drawing productivities of the components from certain predefined probability distributions (Jones, 2005; Growiec, 2008a,b, 2013).

Based on this set of assumptions, we obtain surprisingly general results regarding the implied cross-sectional distributions of technological productivity. Namely, drawing from the Central Limit Theorem and the Extreme Value Theory, we find that if the number of components of a technology, $n$, is sufficiently large, these distributions should be well approximated either by:

(i) the lognormal distribution — in the case of unitary elasticity of substitution between the components ($\sigma = 1$ as in Kremer, 1993);
(ii) the Weibull distribution — in the case of perfect complementarity between the components (the "weakest link" assumption, $\sigma = 0$ as in Growiec, 2013),
(iii) the Gaussian distribution — in the (empirically very unlikely) case of perfect substitutability between the components ($\sigma \to \infty$),
(iv) a novel "CES/Normal" distribution — in any intermediate CES case, parametrized by the elasticity of substitution between the components ($\sigma > 0, \sigma \neq 1$).

We proceed to investigate the properties of the right tail of the "CES/Normal" distribution. Computing its Pareto as well as Weibull tail index confirms that, if technology components are gross complements but are not perfectly complementary ($\sigma \in (0, 1)$), the tail of this distribution decays faster than the tail of any Pareto distribution (i.e., it does not follow a power law) but slower than the tail of any Weibull distribution.

This tail result is interesting because although the prevalence of fat-tailed distributions has been documented and thoroughly discussed for firm sizes, along with a wide array of other

☆ The views expressed herein belong to the author and have not been endorsed by Narodowy Bank Polski. The author thanks Raouf Boucekkine, Jacek Suda, an anonymous Referee of the *NBP Working Papers* series, and participants of the EcoMod (Prague, 2013) and "Structural Change, Dynamics, and Economic Growth" (Livorno, 2013) conferences for their helpful comments and suggestions. Financial support from Warsaw School of Economics (project KAE/BMN/27/14) is gratefully acknowledged. All errors belong to the author.

* Correspondence to: Katedra Ekonomii Ilościowej SGH, al. Niepodległości 162, 02-554 Warszawa, Poland. Fax: +48 22 564 9322.
*E-mail address:* jakub.growiec@sgh.waw.pl.

http://dx.doi.org/10.1016/j.jmateco.2015.06.004
0304-4068/

phenomena in economics and finance,[1] clearly *most* economic variables do not have this property.[2] The distribution of technological productivity, with which we deal here and which has not (to our knowledge) been studied in the empirical literature so far, is in turn one of the important primitives for the firm size distribution. Hence our finding that, from the theoretical point of view, technological productivity distributions should *not* be expected to be fat-tailed, indicates that the apparent emergence of power law tails in firm size distributions[3] must be driven by other phenomena, such as, e.g., endogenous technology choice (Jones, 2005; Growiec, 2008a,b), resource misallocation (Hsieh and Klenow, 2009; Jones, 2011), or aggregation across multi-product firms (Fu et al., 2005; Growiec et al., 2008).

Our aforementioned theoretical contribution to the literature is supplemented with a series of numerical simulations, allowing us to approximate the rate of convergence of the true technological productivity distribution to the theoretical limit with $n$. We identify this rate to be fast in the body of the distribution but slow in the tails which capture rare events. We also numerically assess the dependence of the limiting "CES/Normal" distribution on the degree of complementarity between the technology components, $\sigma$.

Potential empirical applications of the theoretical result, reaching beyond the scope of the current paper, include providing answers to the following research questions:

- Does the "CES/Normal" distribution derived here (Eq. (15)) fit the data on firm sizes, sales, R&D spending, etc.? What is the implied value of $\sigma$?
- Do industries differ in terms of their technology complexity as captured by $n$?
- Do industries differ in terms of the complementarity of technology components as captured by $\sigma$?
- How do firms' optimal technology choices and production function aggregation enter the picture? What are the implications for the shape of the aggregate production function?

The remainder of this paper is structured as follows. Section 2 sets up the model and provides the principal analytical results. Section 3 presents the numerical results. Section 4 concludes.

## 2. The model

### 2.1. Distributions of complex technologies

The point of departure of the current model is the assumption that technologies, invented within the R&D process, are inherently complex and consist of a large number of complementary components. Formally, this can be written down in the following way.

**Assumption 1.** The R&D process determines the productivity of any newly invented technology $Y$ as a constant elasticity of substitution (CES) aggregate over $n \in \mathbb{N}$ independent draws $X_i$, $i = 1, \ldots, n$, from the elementary idea distribution $\mathcal{F}$:

$$Y = \begin{cases} \min\{X_i\}_{i=1}^n, & \theta = -\infty, \\ \left(\frac{1}{n}\sum_{i=1}^n X_i^\theta\right)^{1/\theta}, & \theta \in (-\infty, 0) \cup (0, 1], \\ \prod_{i=1}^n X_i^{1/n}, & \theta = 0. \end{cases} \tag{1}$$

The elementary distribution $\mathcal{F}$ is assumed to have a positive density on $[w, v]$ and zero density otherwise (where $w \geq 0$ and $v > w$ can be infinite). For the case $\theta = -\infty$, it is also assumed to satisfy the condition of a regularly varying lower tail (Leadbetter et al., 1983):

$$\lim_{p \to 0+} \frac{\mathcal{F}(w + px)}{\mathcal{F}(w + p)} = x^\alpha \tag{2}$$

for all $x > 0$ and a certain $\alpha > 0$. For the cases $\theta \in (-\infty, 0) \cup (0, 1]$, it is assumed that $EX_i^\theta < \infty$ and $D^2(X_i^\theta) < \infty$. For the case $\theta = 0$, it is assumed that $E \ln X_i < \infty$ and $D^2(\ln X_i) < \infty$.

The parameter $n$ in the above assumption captures the number of constituent components of any given (composite) technology, and thus measures the complexity of any state-of-the-art technology. The substitutability parameter $\theta$ is related to the elasticity of substitution $\sigma$ via $\theta = \frac{\sigma-1}{\sigma}$, or $\sigma = \frac{1}{1-\theta}$. The case $\theta < 0$ captures the case where the components of technologies are gross complements ($\sigma \in [0, 1)$), whereas $\theta \in (0, 1]$ implies that they are gross substitutes ($\sigma > 1$).

It should be noted at this point that, as argued repeatedly by Kremer (1993), Jones (2011) and Growiec (2013), the gross complementarity case is much more likely to provide an adequate description of real-world production processes than the gross subsitutability case. The example of the explosion of the space shuttle *Challenger* due to a failure of an inexpensive O-ring, put forward by Kremer (1993), is perhaps the best possible illustration of the potentially complementary character of components of complex technologies.

More precisely, the *minimum* case (a Leontief function) reflects the extreme case where technology components are *perfectly* complementary, and thus the actual productivity of a complex idea is determined by the productivity of its "weakest link" (or "bottleneck"). This case was assumed in the related contribution by Growiec (2013). Although likely, this case need not hold exactly in reality, since certain deficiencies of design can often be covered by advantages in different respects. The more general CES case captures exactly this possibility (see also Klump et al., 2012).

The limiting Cobb–Douglas case ($\theta = 0$) is the threshold case delineating gross complementarity from gross substitutability. As shown by Kremer (1993), this case is already quite illustrative of effects of complementarity between components of technologies.

Although technical in nature, restriction (2) imposed on elementary probability distributions $\mathcal{F}$ can also be interpreted in economic terms. First, the support of the distribution must be bounded from below by $w$, which means researchers are not allowed to draw infinitely "bad" technologies (zero is a natural lower bound). This rules out distributions defined on the whole $\mathbb{R}$ such as the Gaussian. Second, the pdf of the distribution $\mathcal{F}$ cannot increase smoothly from zero at $w$; there must be a jump. This means that the probability of getting a draw which is "as bad as it gets" cannot be negligible, and this rules out a few more candidate distributions such as the lognormal or the Fréchet. Third, the lowest possible value of the random variable cannot be an isolated atom, which rules out all discrete distributions such as the two-point distribution, the binomial, negative binomial, and Poisson. Yet, the set of distributions satisfying (2) is still reasonably

[1] Including firm sales, firms' R&D spending, asset returns, and city sizes (Sutton, 1998; Gabaix, 1999; Axtell, 2001; Eeckhout, 2004; Clauset et al., 2009; Gabaix, 2009).

[2] In line with the ubiquitous assumption of Gaussian error terms in econometrics.

[3] The existence of power-law tails in empirical distributions has also been contested by some authors, see e.g. Stanley et al. (1995) or Bee et al. (2013).