Decoupling the CES distribution circle with quality and beyond: equilibrium distributions and the CES-Logit nexus

Simon P. Anderson and André de Palma*

Revised May 2019

Abstract

We show for CES demands with heterogeneous productivities that profit, revenue, and output distributions lie in the same closed power-family as the productivity distribution (e.g., the "Pareto circle"). The price distribution lies in the inverse power-family. Equilibrium distribution shapes are linked by linear relations between their density elasticities. Introducing product quality decouples the CES circle, and reconciles Pareto price and Pareto sales revenue distributions. We use discrete choice underpinnings to find variable mark-ups for a more flexible demand formulation bridging CES to Logit and beyond. For logit demand, exponential (resp. normal) quality-cost distributions generate Pareto (log-normal) economic size distributions. (99 words)

JEL Classification: L13, F12

Keywords: CES; monopolistic competition; quality; closed power-family distributions; Pareto, power, generalized (log-)normal; Logit, mark-ups, Box-Cox

*Department of Economics, University of Virginia, USA, sa9w@virginia.edu; CES-University Paris-Saclay, FRANCE. andre.depalma@ens-cachan.fr. The first author gratefully acknowledges research funding from the NSF; the second author gratefully acknowledges research funding from the ANR for the Projet Elitisme. We thank Kerem Cosar, Julien Monardo, Maxim Engers, Isabelle Mejean, Farid Toubal, James Harrigan, Ariell Resheff, for valuable comments, and seminar participants at Hebrew University, Melbourne University, Stockholm University, KU Leuven, Laval, Vrij Universiteit Amsterdam, CUHK, Hitotsubashi, University of Tokyo, Monash, ANU, and Paris Dauphine. We thank particularly Estelle Cantillon and two anonymous referees. This is a revised version of part of CEPR DP 10748. Some of the results in Section 2 are in Mrázová, Neary, and Parenti's CEPR 12044. We would like to thank Eric Verhoogen for discussion on the potential validation of his model.

1 Introduction

Distributions have been studied for a long time in social sciences by geographers and others, and have recently attracted more interest in economics (Axtell, 2001, Anderson, 2006, Gabaix, 2016). Indeed, interest in international trade in sales revenue distributions and in the CES monopolistic competition model has exploded in the empirical literature over the last decade (following Melitz, 2003). The CES monopolistic competition model and a particular sales revenue distribution (the most commonly used ones are the Pareto and log-normal) will tie down the shapes of the other distributions, such as profit, price, and output distributions. However, little research has investigated the other implied distributions. Here we develop equilibrium connections between *distributions* in the paradigm CES demand formulation (which we extend in two major ways) under monopolistic competition. For example, can a log-normal distribution of unit costs be consistent with a log-normal distribution of firms' outputs? Can a Pareto distribution for profit be consistent with a Pareto distribution for firm prices? The answer is affirmative for the first one: in fact, if output is log-normally distributed, then unit costs *must* be log-normally distributed: and so must equilibrium prices. But (for the second question) a Pareto profit distribution function is only consistent with a *power* distribution of prices. However, when we extend the CES for quality heterogeneity we can indeed render Pareto distributions for both.

The CES representative consumer model is widely used in economics in conjunction with monopolistic competition.[1] A flurry of recent contributions deploy the CES and variants thereof (e.g., Dhingra and Morrow, 2017, Zhelobodko, Kokovin, Parenti, and Thisse, 2012, Bertoletti

[1] See the original book on monopolistic competition by Chamberlin (1933).

and Etro, 2017, etc.). The current most intensive use of the model is in International Trade, where it is at the heart of empirical estimation.[2] It is used as a theoretical component in the New Economic Geography and Urban Economics, it is the linchpin of Endogenous Growth Theory, Keynesian underpinnings in Macro, and Industrial Organization. The convenience of the model stems from its analytic manipulability. The CES monopolistic competition model delivers equilibrium mark-ups proportional to marginal costs, and so delivers market power (imperfect competition) in a simple way without complex market interaction. The standard models in this vein (following Melitz, 2003) assume that firms' unit production costs are heterogeneous.

However, when we apply this model to distributions, if one distribution (such as profit) is a Pareto (1965, 1896) distribution, then the distributions of all the "primary" economic variables (productivity, revenue, profit, output, and price reciprocal) lie in the Pareto class. This we call the "Pareto circle". More generally, we establish the *CES circle* by characterizing distribution families that are closed under the positive power transformations that the CES entails between the primary economic variables. These distributions we call *Closed Power-Families* - CPF henceforth. The implication is that if any one of the distributions is in the family, then they all are. Previous authors have derived special cases of this relation between productivity and sales revenue for particular distributions: Helpman, Melitz, and Yeaple (2004) cover Pareto, while Head, Mayer, and Thoenig (2014) treat the log-normal distribution. We show that the Pareto and log-normal distributions constitute closed power-families and we show that these authors' results cover all power-closed distribution families as well as the other economic variables.

On the other hand, other standard distributions are *not* closed under power transformations.

[2] Although note that Fajgelbaum, Grossman, and Helpman (2011) take a nested multinomial logit approach.

If the productivity distribution is normal or Laplace, the output and revenue distributions are not (except under exceptional circumstances for specific parameter matches). But, both these distributions will seat in a broader class of CPF.

Another key characterization result is that if the output, revenue, and productivity distributions form a CPF, then their densities take the same functional form as the *price* density. Indeed, price is determined from the other variables via a negative power transformation for the CES. The relevant distributions for negative power transformations are thus the inverse distributions to the positive power family. Indeed, a distribution and its survival function (e.g. Pareto and power) have the same parametric density function. This property connects their densities. One straightforward consequence of the density representation is that the price-reciprocal distribution is log-normal if the revenue distribution is log-normal.

The power-family analysis enables us to write the relation between the *shapes* of the satellite distributions as simple linear relations of density elasticities and the demand elasticity of the CES. For example, the elasticities of both output and profit densities are linear functions of the elasticity of the density of costs, with the demand elasticity determining the parameters. This analysis extends the insights from the constant elasticity case associated to the Pareto distribution. For instance, it enables us to describe inherited concavity/convexity properties of satellite distributions. It also enables us to derive some testable predictions for the densities, for example the relationship between their respective modes.

The CES circle imposes restrictions on the satellite distributions. The (ubiquitous) CES demand side can be retained while decoupling the CES circle by introducing qualities. These we introduce in the same way as do Baldwin and Harrigan (2011) and Feenstra and Romalis

(2014).[3] Doing so delivers two fundamental drivers of equilibrium distributions (instead of just one) – the cost distribution and the quality/cost one. Even if one distribution is Pareto, others can take other forms. Most notably, the output distribution depends on the cost distribution (as before) but now *also* on the quality/cost distribution and so there are three groups of distributions involved. This we call the *decoupling* of the CES circle. Paradoxically, perhaps, introducing quality also enables a *recoupling* of distributions. As noted above, without quality, a Pareto revenue distribution must induce a power distribution for prices. But, with a quality distribution too, and the appropriate (sufficiently positive) relation between quality and cost, the price distribution can be Pareto too. However, the mark-ups in percentage terms remain constant across products due to the CES formulation, jarring with several empirical studies (e.g., Atkin et al., 2015). This property spurs us to develop a broader demand formulation which allows for variable mark-ups.

The CES demand system is usually derived from a representative consumer's utility function, although individual consumers typically consume few variants in practice. The demand system though is also consistent with a population of heterogeneous consumers who make discrete choices across products. This disaggregation approach enables us to formulate a generalization which picks up the CES as a special case and links the CES to the logit model of monopolistic competition (which we propose and develop here) via a structural demand parameterization using an additional parameter, b. Our model delivers clean comparative static results and profit quasi-concavity ensuring a unique maximum. The case $b = 0$ corresponds to the CES, while $b = 1$ is Logit. For $b \in (0, 1)$, percentage mark-ups are increasing in b while absolute

[3] These authors assume a Pareto distribution for productivities for their empirical work. They do not pursue the implications of the properties of the various distributions and how they are linked, which is our goal here.

mark-ups are decreasing. Not only can we book-end the model with CES and Logit, but we can also go beyond them. In particular, the (well-behaved) case $b < 0$ delivers the property that even percentage mark-ups rise with cost, and with quality too as long as the unit cost function is sufficiently inelastic with respect to quality. Such properties are consistent with empirical findings of Atkin et al. (2015) that mark-ups are increasing with quality, and larger firms have higher qualities. We also deliver and develop new distribution results for Logit: a normal distribution for quality-costs implies a log-normal distribution for firm size in terms of both output and profit, while an exponential function delivers Pareto distributions. The log-normal and Pareto distributions have figured prominently in empirical work on size distribution of firms.

A wide-ranging paper by Mrázová, Neary, and Parenti (2017) also links distributions and demand systems, focussing on Constant Revenue Elasticity of Marginal Revenue (CREMR) and Constant Elasticity of Marginal Revenue cases; the intersection of the two is the CES form. They show various "self-reflection" properties; most notable for our study are their 3-way equivalence properties, whereby any two properties imply the third. In particular, they link (i) Pareto productivity and sales revenue distributions to CREMR; (ii) log-normal productivity and sales revenue distributions to CES; (iii) Pareto productivity and output distributions to CEMR; and (iv) log-normal productivity and output distributions to a case of CEMR. Their paper is also noteworthy for analyzing how Kullback-Leibler divergence measures can be used for comparing predicted to observed distributions, concluding that the choice of Pareto or Log-Normal distribution is less important than the demand form.

Effectively, they add important additive constant parameter to the early part of our analysis.

Their philosophical backdrop follows Anderson and de Palma (2015) in studying how endogenous economic distributions are related to each other and to primitive distributions through the core economic monopolistic competition model. Methodologically, noteworthy is our use of our family density equivalence result as a device to uncover the tightest Closed Power-Family encompassing the Normal, and we give parameter relations that must hold to be consistent with the CES model, while MNP concentrate on log-normal and Pareto distribution circles (as noted above). For example, we define the parameterization of the A-family which includes the Normal (for a Normal alone transforms into other distribution types) within the tightest possible class; and we seat the log-normal within a broader class too (the log-A family). We also extend to quality, and our generalization (which enables us to pick up the Logit) does not fall into MNP's CREMR or CEMR families.

In the sequel, we first develop the analysis for cost heterogeneity alone and show how the productivity distribution delivers the Pareto (and power) circle. In Section 3, we then describe the general CES circle and explore further power families of distributions. We next show in Section 4 how equilibrium densities are tied together by simple linear relations between their elasticities. Section 5 allows for quality too in order to decouple the CES circle. Section 6 turns to a disaggregate micro-foundation of the CES representative consumer to generalize the demand model, and tie together CES and Logit special cases. A final section gives some concluding remarks.

2 Standard CES model and the Pareto circle

2.1 Basic CES model with cost heterogeneity

We start with the standard CES monopolistic competition model with heterogeneity only in firms' productivities (the reciprocal of unit production costs). This is the basic Melitz (2003) approach. We show how all economic distributions (output, profit, revenue, and price reciprocal) are tied down by the productivity distribution, and we later show (in Section 3.3) that this distributional relationship defines a closed power-family to which all these satellite distributions belong.

Several forms of CES representative consumer utility functions are prevalent in the literature. We nest these into one embracing form. The CES representative consumer involves a sub-utility functional for the differentiated product $\chi = \left(\int_{\Omega} q(\omega)^{\rho} d\omega\right)^{1/\rho}$, where the q's are quantities consumed of the differentiated variants, and $\rho \in (0,1)$: variants are perfect substitutes for $\rho = 1$, demands are independent for $\rho \to 0$. The elasticity of substitution is $\sigma = \frac{1}{1-\rho} \in (1,\infty)$. This statistic is often reported in empirical studies: $-\sigma$ is also the (constant) elasticity of the CES demand function. The important property of the CES for what follows is that the demand elasticity is constant (and elastic).

The individual variants are denoted by ω, and each is produced by a separate firm; the set of variants is denoted by Ω. Common forms of representative consumer formulation are:

(i) Melitz (2003) model (see also Dinghra and Morrow, 2017), where $U = \chi$ so there is only one sector;

(ii) the classic Dixit-Stiglitz (1977) case much used in earlier trade theory (e.g., Helpman

and Krugman, 1985), $U = \chi q_0^{\eta}$ with $\eta > 0$, where q_0 is consumption in a numeraire sector;

(iii) $U = \ln \chi + q_0$, which is a quasi-linear form (with no income effects) and so constitutes a partial equilibrium approach (see Anderson and de Palma, 2000, and Nocke and Yeaple, 2008).

The first two formulations have unit income elasticities of demand; hence their popularity in Trade models. Utility is maximized under the budget constraint $\int_{\Omega} q(\omega) p(\omega) d\omega + q_0 \leq I$, where $p(\omega)$ is the price of variant ω. We need $I > 1$ for case (iii), or else it reverts to case (i) because all income is spent on the differentiated variants.

Let the (constant) unit production costs of variant ω be $c(\omega)$, and let $F_C(c)$ denote the cumulative distribution function of these costs. Let $f_C(c)$ denote the probability density of unit production costs, and assume that it is continuously differentiable.[4]

The next results are quite standard. For a given set of prices and a set Ω of firms (with total mass $M = \|\Omega\|$), Firm i's demand is:

$$h(p_i) = \frac{\Xi(I)}{p_i} \frac{p_i^{\frac{\rho}{\rho-1}}}{\int_{\omega \in \Omega} p(\omega)^{\frac{\rho}{\rho-1}} d\omega}, \tag{1}$$

where $\Xi(I)$ is the total amount spent on the differentiated commodity. This is $\Xi(I) = I$ for case (i), $\Xi(I) = \frac{I}{1+\eta}$ for case (ii) (which clearly nests case (i) for $\eta = 0$); and $\Xi(I) = 1$ for case (iii). The denominator in (1) (assumed bounded) represents the aggregate impact of firms' actions on individual demand: under monopolistic competition, each firm's action has no effect on this statistic. Notice that the CES demand system exhibits the IIA (Independence from Irrelevant Alternatives) property that the ratio of the demands for any two products is independent of

[4] Exceptions will be flagged when relevant.

the price of any other product.[5] The analysis of Section 6 indicates why the CES shares this property with Logit discrete choice models.

Firm i's profit maximizing price solves $\max_{p_i} \left\{ \frac{(p_i - c_i)}{p_i} p_i^{\frac{\rho}{\rho-1}} \right\}$, so equilibrium $p_i = \frac{c_i}{\rho}$, and the equilibrium Lerner index is

$$l_i \equiv \frac{p_i - c_i}{p_i} = (1 - \rho). \tag{2}$$

Given such pricing, Firm i's equilibrium output is (from (1))

$$y_i = \rho \Xi (I) \frac{c_i^{\frac{1}{\rho-1}}}{D_C}, \tag{3}$$

where $D_C = M \int c(u)^{\frac{\rho}{\rho-1}} f_C(u)\, du$, and $f_C(.)$ is the density of unit production costs, which we assume ensures D_C is bounded.

Firm i's equilibrium profit, π_i, is proportional to its sales revenue, $r_i = p_i y_i$, so that

$$r_i = \Xi (I) \frac{c_i^{\frac{\rho}{\rho-1}}}{D_C}, \tag{4}$$

and

$$\pi_i = (1 - \rho) r_i. \tag{5}$$

Hence equilibrium output is elastic with respect to cost, with elasticity $\frac{1}{\rho-1} = -\sigma < -1$, which is the same as the *elasticity of demand* for the CES. Sales revenue (and also profit) has elasticity $\frac{\rho}{\rho-1} = 1 - \sigma < 0$, which is smaller in absolute terms. Doubling cost cuts output by more than half, while profit goes down proportionately less (because of the price increase).

[5] That is, from (1), $h(p_i)/h(p_j) = (p_i/p_j)^{1/(\rho-1)}$ is independent of p_k, $k \neq i, j$.

2.2 Pareto and power distributions

Part of the folklore for the CES is that "everything is Pareto", although we cannot cite a definitive reference. Helpman, Melitz, and Yeaple (2004, p. 304) showed that the *revenue* distribution is Pareto if the *productivity* distribution is Pareto for the CES, and Mrázová, Neary, and Parenti (2017) showed the same relation for revenue and output distributions. Here, for the record, we make the statement precise by documenting the *CES-Pareto circle* for all five *primary distributions* (i.e., the ones above, which we refer to as primary because they will all be seen to related by a simple transformation). We simultaneously make the analogous statement for the power distribution.

Define $\hat{c} \equiv 1/c$ as productivity, and suppose that the distribution of productivities is Pareto,

$$F_{\hat{C}}(\hat{c}) = 1 - \left(\frac{\underline{\hat{c}}}{\hat{c}}\right)^{\alpha_{\hat{c}}} \text{ for } \hat{c} \geq \underline{\hat{c}}, \tag{6}$$

where $\alpha_{\hat{c}} > 0$.[6] Now, (4) implies $r = k\hat{c}^{\theta}$, where θ takes the value $\frac{\rho}{1-\rho} > 0$. Therefore

$$F_R(r) = 1 - \left(\frac{\underline{r}}{r}\right)^{\alpha_r} \text{ for } r \geq \underline{r}, \text{ with } \alpha_r = \alpha_{\hat{c}}\frac{1-\rho}{\rho}.$$

From (5), this tail parameter value is also α_π, with $F_\Pi(\pi) = 1 - \left(\frac{\underline{\pi}}{\pi}\right)^{\alpha_\pi}$ for $\pi \geq \underline{\pi}$.

Because from (3) equilibrium output has productivity elasticity $\frac{1}{1-\rho}$, the equilibrium output has a Pareto distribution with tail parameter

$$\alpha_y = (1-\rho)\,\alpha_{\hat{c}} = \alpha_r \rho. \tag{7}$$

[6] The $\underline{\hat{c}}$ value might be endogenously determined by fixed cost, as in Melitz (2003).

It is readily apparent that the price reciprocal is also Pareto, thus establishing that the primary distributions are all Pareto if any one is.

The cost distribution associated to the Pareto productivity distribution (6) is the power cost distribution:

$$F_C(c) = \left(\frac{c}{\underline{c}}\right)^{\alpha_{\hat{c}}} \text{ for } c \in [0, \underline{c}], \tag{8}$$

and the price distribution has the same tail parameter, so:

$$F_P(p) = \left(\frac{p}{\underline{p}}\right)^{\alpha_{\hat{c}}} \text{ for } p \in \left[0, \underline{p}\right], \text{ where } \underline{p} = \underline{c}/\rho.$$

Following the analogous steps above shows that the five primary distributions are Power distributions, with the same relation as (7) holding for their power values. Furthermore, the cost and price distributions are Pareto, with tail parameter $\alpha_{\hat{c}}$. To summarize:[7]

Proposition 1 *(CES Pareto / Power circle) (a) If any one of the five primary distributions (profit, revenue, output, price reciprocal, or productivity) is Pareto (resp. Power), then they all are Pareto (resp. Power). (b) Price and cost are Power distributed if the five primary distributions are Pareto, and are Pareto distributed if they are Power. (c) Tail/power parameters are related by (7):* $\alpha_y = (1-\rho)\,\alpha_{\hat{c}} = \alpha_r \rho$; $\alpha_r = \alpha_\pi$; $\alpha_p = \alpha_{\hat{c}}$.

The Pareto circle says that all five (primary) distributions in (a) are Pareto if any one is. Part (b) indicates that the survivor functions for price and cost also take the Pareto form. The positive transformation sends any Pareto distribution to another Pareto, while a negative transformation sends a Pareto to the inverse circle, which is the Power one. The Power circle

[7] Part (c) is discussed further in Section 4.

is analogous, mutatis mutandis.[8]

3 Distribution transformations, Closed Power Families, and the CES-circle

Multiplicative and positive power transformations relate profit, revenue, output, price reciprocal $(1/p)$, and productivity (the cost reciprocal, $1/c$) in the CES model. The corresponding primary distributions are related with the help of the following straightforward technical result, which tells us how distributions are modified by (positive) *multiplicative* and *power transformations*.[9] We shall also be interested in cost and price distributions, which are related to the others by negative power transformations. Let U be defined on a bounded support.

Lemma 1 *(Transformation) Let $F_U(u)$ be the CDF of a random variable U. Then, the CDF of $F_V(v)$ for the multiplicative transformation ($V = kU$, $k > 0$), positive power transformation ($V = kU^\theta$, $k > 0$, $\theta > 0$), and negative power transformations ($V = kU^\theta$, $k > 0$, $\theta < 0$) are:*

(a) (Multiplicative) $F_V(v) = F_U\left(\frac{v}{k}\right)$;

(b) (Positive power) $F_V(v) = F_U\left[\left(\frac{v}{k}\right)^{\frac{1}{\theta}}\right]$;

(c) (Negative power) $F_V(v) = 1 - F_U\left[\left(\frac{v}{k}\right)^{\frac{1}{\theta}}\right]$.

This coheres with the earlier Pareto and power distribution analysis. For example, power distributions beget power distributions under positive power transforms. They beget Pareto distributions under negative power transforms and conversely: Pareto distributions beget power distributions under negative power transforms. The Lemma defines *multiplicative and power*

[8] As we show below, the log-normal also delivers a CES circle for the five primary distributions.

[9] For example, for case (b) below, $F_V(v) = \Pr(V < v) = \Pr\left(kU^\theta < v\right) = \Pr\left(U < \frac{v}{k}^{\frac{1}{\theta}}\right) = F_U\left[\left(\frac{v}{k}\right)^{\frac{1}{\theta}}\right]$.

relations between distributions, which we immediately engage to characterize the CES distribution circle:

Proposition 2 *(CES distribution relation) (a) The distributions of profit and revenue are multiplicatively related, as are the distributions of price and cost. (b) The primary distributions (profit, revenue, output, price reciprocal, and productivity) are positive-power related. (c) The distributions of price and cost are negative-power related to the distributions of profit, revenue and output (and conversely).*

Proof. From (5), profit is a positive fraction of revenue. By Lemma 1(a), revenue and profit distributions are related by multiplicative transformations. Likewise, cost is a positive fraction of equilibrium price (see (2)), so the same property attains. From (3), equilibrium output, y_i, is related to the cost reciprocal, $1/c_i$, by a positive power and a positive factor, and similarly for the price reciprocal, $1/p_i$. Profit and revenue distributions are related to the cost reciprocal and hence to the output distribution by a positive-power transformation. Hence the second statement follows from Lemma 1(b). The last statement follows from 1(c) because costs and prices are related to output and to revenue and profit by negative-power transformations. ■

3.1 The CES Circle

We here extend the Pareto/power circle to provide some distributions that form what we call *Closed Power-Families (CPF)* under the CES transformation, which transformation for the primary distributions (as we have seen) involves taking positive powers. In such a closed power-family, the satellite distributions retain the same functional form under the positive

power transformation, so that profit, revenue, output, productivity (cost reciprocal), and price reciprocal remain in a CES-circle. Because we want the results to hold for any CES specification (any value of $\rho \in (0,1)$), we want the family to entertain any positive power transformation.[10]

We first provide a formal definition of the concept of a CPF. If $\mathcal{F}(.)$ is to be such a family, then, setting $v = ku^{\theta}$ with $k > 0$ and $\theta > 0$ must imply that $F\left(\left(\frac{v}{k}\right)^{1/\theta}\right) \in \mathcal{F}(.)$ for *any* $\theta > 0$.[11] Then

$$F_V(v) = F_U\left[\left(\frac{v}{k}\right)^{1/\theta}\right]$$

(with $k > 0$ and $\theta > 0$) must hold for any pair of distributions of U and V in the family (and similarly for $\theta < 0$, see Lemma1 (c)).

Differentiating this identity leads us to a density relation that any two densities in the family must satisfy:

$$f_V(v) \propto v^{\left(\frac{1}{\theta}-1\right)} f_U\left(\left(\frac{v}{k}\right)^{1/\theta}\right). \tag{9}$$

This empirically-testable density expression is the *key* to finding candidate power-closed families (and eliminating others): if the relation does not hold, then the CES model as given is empirically invalid.[12] Indeed, we have been working till now with CDFs, and for the Pareto and Power distributions they are the appropriate tool. But if we consider other distributions (like Normal, or Log-Normal) there is no analytical form for the CDF, and *a fortiori* for their power transformations. A more tractable way to relate such distributions is through their densities, and this approach enables us to show that the log-normal constitutes a CPF, while the Normal needs to be extended as an element of a larger class.

[10] Recall that the relevant θ can take any value in $(0,\infty)$ because $\rho \in (0,1)$: see Section 2.
[11] Here $u \in [\underline{u}, \bar{u}]$ so $v \in \left[k\underline{u}^{\theta}, k\bar{u}^{\theta}\right]$ for $\theta > 0$ and $v \in \left[k\bar{u}^{\theta}, k\underline{u}^{\theta}\right]$ for $\theta < 0$.
[12] Notice that power and Pareto distributions both deliver power densities.

3.1.1 CPF densities

Consider now the shape of the densities of the inverse distributions: the ones of interest are cost and price (which vary inversely with the primary variables.) This implies that if a distribution U is a closed power-family then U^{θ} belongs to this family for $\theta > 0$ and U^{θ} belongs to the inverse family for $\theta < 0$. From Lemma 1(c), for $\theta < 0$ we have the survivor function $F_V(v) = 1 - F_U\left[\left(\frac{v}{k}\right)^{\frac{1}{\theta}}\right]$. Differentiating, we get exactly the same expression (9) as for the case of a positive power transformation (i.e., $F_V(v) = F_U\left[\left(\frac{v}{k}\right)^{\frac{1}{\theta}}\right]$ with $\theta > 0$). Therefore we can write the following Corollary of Lemma 1 and Proposition 2:

Corollary 1 *(CES densities) The densities of profit, revenue, output, price, and cost for a closed power-family satisfy (9).*

We now turn to some closed power-families. While standard distributions such as the Normal cannot represent all the primary distributions, we show how to extend such distributions to seat them within a corresponding circle. However, the Log-Normal is already a closed power-family.

3.2 Log-Normal

The *log-normal distribution* also delivers a CES circle. Head, Mayer, and Thoenig (2014) showed that a log-normal distribution for productivity delivers log-normal revenues for the CES.[13] We show how this result extends to encompass all the primary distributions – as well as the distributions of their reciprocals.

[13] These authors look at the empirical evidence for choosing either log-normal or Pareto distributions for firm size: Nigai (2017) splices the two (with the right tail Pareto) to get the best fit.

To see this property, consider the basic log-normal density

$$f_U(u) \propto \frac{1}{u} \exp\left(-\left(\frac{\ln u - \xi}{\sigma}\right)^2\right), \qquad u > 0. \tag{10}$$

Making the change of variables $v = ku^{1/\theta}$, as befits the basic power transform fundamental to the CES circle, then gives another log-normal by simply adjusting ξ and σ appropriately, and this is true for θ positive or negative. The immediate implication is that the revenue and productivity distributions are log-normal if one is (as shown by Head et al., 2014). Furthermore, so are profit and output (the output result was shown by Mrázová, Neary, and Parenti, 2017), and so are price and cost densities (which involve the survivor function distributions because of the flip from high prices associated with low profits, etc.). Empirically, the log-normal has shown up for revenue distributions in Head, Mayer, and Thoenig (2014) among others. To summarize:

Corollary 2 *(Log-Normal densities) The densities of profit, revenue, output, price, and cost all have the form (10) if any one has.*

3.3 Other CPF distributions for CES

3.3.1 A-family

Recent empirical work has not supported power distributions for prices (and hence, for the CES, for costs either). Two eminent studies with big data on prices have recently appeared: Kaplan and Menzio (2015) and Hitsch, Hortacsu, Lin (2017). While the latter do not directly estimate the shape of price distributions, Kaplan and Menzio (2015) tend to support symmetric bell-shaped densities (although a huge caveat is that they deal with prices across a range of

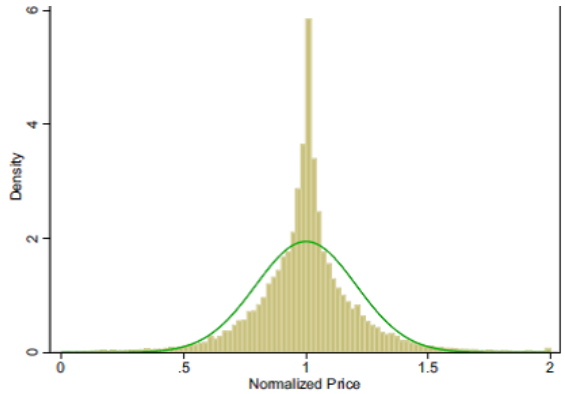

Figure 1: Price density from Kaplan and Menzio (2015)

retail outlets and for a broad gamut of products, rather than the specific industries we have in mind, although they do find similar patterns for different product groupings) The following Figure reproduces their Figure 2a (Kaplan and Menzio, 2015, p. 1174) with the empirical distribution approximated by a Normal (green curve):

Notice that this empirical distribution looks closer to a Laplace than a Normal: in what follows we consider a class of densities that includes both as special cases.[14]

On the other hand, several studies suggest (e.g., Head, Mayer, and Thoenig, 2014; Cabral and Mata, 2003) a Pareto or a log-normal for sales revenue or profit. As we have shown above, a Pareto profit distribution implies a power price distribution under CES. We now entertain a reverse question, and ask what types of (revenue or output) distribution are delivered by normal or log-normal price densities (and we cover log-normal revenues along the way). In what follows, it is easier to work directly with the density, with parameters determined so that

[14]From Kaplan and Menzio (2015): "First, for all definitions of a good, the price distribution has a unique mode that is very close to its mean. Second, for all definitions of a good, the price distribution is very close to symmetric. Third, for all definitions of a good, the price distribution has more mass around the mean and has thicker tails than a Normal distribution with the same mean and variance."

the densities do generate distributions over the relevant support. Truncated distributions are handled through writing the appropriate support for the transformed variable.

We start by looking at some distributions that are not closed power-families, in order to guide us towards ones that are. Consider first the Normal, which has density $f_U(u) \propto \exp\left(-\left(\frac{u-\xi}{\sigma}\right)^2\right)$. The corresponding density of V is *not* a Normal (except for specific parameter values noted below), and for two reasons. Applying (9) (and using $dv = k\theta u^{\theta-1}du = k\theta\left(\frac{v}{k}\right)^{\left(1-\frac{1}{\theta}\right)}du$) gives $f_V(v) \propto v^{\left(\frac{1}{\theta}-1\right)}\exp\left(-\sigma^{-2}\left(\left(\frac{v}{k}\right)^{1/\theta}-\xi\right)^2\right)$, so that there is an extra power term before the exponential term, and, second, the term in the square now has a power not equal to 1 (if $\theta \neq 1$).[15]

Similarly, a Laplace density with $f_U(u) \propto \exp\left(-\left|\frac{u-\xi}{\sigma}\right|\right)$ does not constitute a closed power-family. A Fréchet distribution (see Mrázová, Neary, and Parenti, 2017), given by $f_U(u) \propto u^{-\alpha-1}\exp\left(-\frac{(u-\mu)^{-\alpha}}{\sigma}\right)$, is not closed under a power transformation. However, a simple Fréchet distribution (without shift parameter and with $u > 0$) of the form $f_U(u) \propto u^{-\alpha-1}\exp\left(-\frac{u^{-\alpha}}{\sigma}\right)$ does constitute a closed power-family because the same form arises under the power transformation.[16]

If $\theta = 1$, the distributions of all economic variables are the same up to a (multiplicative) scale factor. Given the multiplicative relation for the CES between revenues and profits, and also between productivity and price reciprocal, these pairs of variables follow the same distribution, up to a scale factor.[17]

[15] If $\theta = 1$ then neither of these issues arise, a point we elaborate in the next paragraph.

[16] The same argument applies to the double Weibull (given in the footnote below) without shift parameter and with $x > 0$, whose density has the form $f_U(u) \propto u^{\alpha-1}\exp\left(-\left(\frac{u-\zeta}{\sigma}\right)^{\alpha}\right)$.

[17] There is one other case when θ can be 1, and that is the relationship between revenue (or profit) and productivity when $\rho = 1/2$. Then, from (4), revenue linearly tracks productivity.

The Normal and Laplace distributions can be extended to yield power-closed families, which are parameterized by A below. That is, any value of A defines a power-closed family, which we call an *A-family*: all of the other parameters vary according to the satellite distribution while remaining in the A-family.

The A-family is defined by:

$$f_U(u) \propto u^{\Lambda} \exp\left(-\left|\frac{u^B - \xi}{\sigma}\right|^A\right), \quad u \geq 0, \tag{11}$$

with parameters $\{\Lambda, B, \sigma, \xi; A\}$.[18] Throughout, we set $\sigma > 0$ and we take $u \geq 0$ as relevant for the economic variables under consideration.[19]

This specification encompasses standard distributions such as the power or Pareto (for $A = 0$); the Normal for $\{\Lambda, B; A\} = \{0, 1; 1\}$; the Laplace for $\{\Lambda, B; A\} = \{0, 1; 1\}$; and, for simple Fréchet (with no shift), $\Lambda = -\alpha - 1$, $AB = -\alpha$, and $\xi = 0$. The power transformation then launches satellite distributions in the same A-family. The density (11) is continuous and is differentiable except at $u = \xi^{1/B} \geq 0$ (although it is differentiable if $A = 2$).

Under the CES transformation $v = ku^{\theta}$, we have the V density satisfying

$$f_V(v) \propto v^{\Lambda_V} \exp\left(-\left|\frac{v^{B_V} - \xi_V}{\sigma_V}\right|^A\right), \quad v > 0, \tag{12}$$

where the V subscripts indicate how the parameters change under a positive power transform,

[18] This form is similar to the double Weibull, which is $f_U(u) = \frac{c}{2\sigma}\left|\frac{u-\zeta}{\sigma}\right|^{A-1} \exp\left(-\left|\frac{u-\zeta}{\sigma}\right|^A\right)$. For $A = 1$, we get the Laplace (or the double exponential) density. The double Weibull generalizes the Rayleigh distribution, whose density is $u \exp\left(-u^2\beta\right)$. However, unless $\zeta = 0$, the double Weibull does not constitute a power class. However, introducing a power parameter on the u on the RHS rectifies this.

[19] As noted in Section 3.1, trunctations for this density and the following ones may be required if the CDF is not defined on $(0, \infty)$, or if some moments of the distribution do not exist on $(0, \infty)$, as for the case of the *generalized log-normal* considered below.

and hence the closed density power family dictates $\Lambda_V = -1 + (1+\Lambda)/\theta$, $B_V = B/\theta$, $\xi_V = \xi k^{B/\theta}$, and $\sigma_V = \sigma k^{B/\theta}$. The transformed parameters indicate what to expect empirically when the various distributions are estimated from this family, since we know how parameters are related. The transformation remains in the same CPF.

As we argue below, the same arguments as above apply to price densities, so that we have the result that price densities are in the same closed density power-family too. To illustrate, consider the form (12), and the relation between productivity (1/c) and price, which for CES is the relation $p = 1/\hat{c}\rho$. Hence, the power transform for price is to set $k = 1/\rho$ and $\theta = -1$ (negative power.) Then the price density (denoted by subscript P) is (using $\Lambda_P = -2 - \Lambda$, $B_P = -B$, $\xi_P = \xi\rho^B$, and $\sigma_P = \sigma\rho^B$ in (12)):

$$f_P(p) \propto p^{-2-\Lambda} \exp\left(-\left|\frac{(\rho p)^{-B} - \xi}{\sigma}\right|^A\right), \quad p > 0.$$

Note the price distribution is a power distribution in the special case of a Pareto productivity distribution.[20] Reciprocally for a power productivity distribution: the price distribution is then Pareto. As we show next, the log-normal distribution for productivity begets prices that are also log-normal (along with all the other satellite distributions being log-normal).

3.3.2 Log-A-family

A potentially useful generalization of the log-Laplace and log-Normal that is a closed density power-family (for any A) as an alternative to (11) is to write the *log-A*-family.

[20] Discussed more in the next section.

The *log-A*-family is defined by densities of the form

$$f_U(u) \propto \frac{1}{u} \exp\left(-\left|\frac{\ln u - \xi}{\sigma}\right|^A\right), \qquad x > 0. \tag{13}$$

This distribution has previously appeared in the literature as the *generalized log-normal* (Kleiber, 2014).

Here, when $A = 2$, the density is log-normal; and when $A = 1$ it is log-Laplace. Indeed, for any given A, (13) constitutes a closed $logA$-power-family. This form also nests Pareto, and potentially constitutes a flexible form to estimate, in a manner that extends the log-normal and Pareto. Not only does this allow tighter investigation of the CES circle, where log-normal and Pareto have played prominent roles, but it also can be a useful guide for relating distributions in other contexts.

4 Distribution shape relations

4.1 Constant elasticity productivity distributions

For the Pareto distribution, densities are always decreasing (the distribution function is concave), so profit, revenue, and output densities must decrease. However, the price or cost ones can be increasing. This can be seen from the relation $\alpha_{\hat{c}} = \alpha_r \left(\frac{\rho}{1-\rho}\right)$ (see (7)) and noting that the slope of the density of costs has the sign of $\alpha_{\hat{c}} - 1$ so that the cost density is decreasing if and only if $\alpha_r = \alpha_\pi < \frac{1-\rho}{\rho}$.[21] For example, if the tail parameter is $\alpha_\pi = 1$ (as per Axtell,

[21] Recall that $\alpha_{\hat{c}}$ is the elasticity of the cost distribution and must be constant to deliver the Pareto distributions for the primary variables. Equivalently, $\alpha_{\hat{c}}$ is the tail parameter of the (Pareto) productivity distribution.

2001) then the cost density is decreasing if $\rho < 1/2$, i.e. far from close substitutes.[22] And if the "80/20" rule holds, then $\alpha_\pi = 1.16$, with corresponding value $\rho < 0.46$ for decreasing cost density.[23]

4.2 General density analysis

While the Pareto imposes decreasing primary densities, this property is not true generally for other distributions. We next show the connections between the densities of the satellite distributions by determining the inheritance properties of distributions under the (positive or negative) power transformation. That is, we determine the elasticity properties of the pdfs that characterize the relations between distributions for the CES model. We consider the relations between cost, output, and revenue distributions (the CES involves power transformations between these variables).

Lemma 2 *(Inheritance) Let $F_U(u)$ be the CDF of a random variable U and let $V = kU^\theta$ with $k > 0$ and $\theta \neq 0$. Let $\eta_{f_U}(v)$ be the elasticity of $f_U(u)$ with $u = \left(\frac{v}{k}\right)^{1/\theta}$. Then*

$$\eta_{f_V}(v) = \frac{1}{\theta}\left\{\eta_{f_U}(v) + (1-\theta)\right\}. \tag{14}$$

Proposition 3 *Hence, if $\eta_{f_U}(v) > -1$, then $\eta_{f_V}(v) > -1$ iff $\theta > 0$; if $\eta_{f_U}(v) < -1$, then*

[22] Axtell (2001) estimates the sales revenue tail parameter as $\alpha_r = .994$ (and recall that $\alpha_\pi = \alpha_r$ for the CES). He estimates the tail parameter for firm size by employee numbers as 1.059, which one might take as a proxy for α_y. Both are estimated for 1997 US Census Bureau data. Intriguingly, (7) then suggests that ρ is close to 1 (indeed, larger!), at least for the aggregate data. However, empirical estimates of ρ from Broda and Weinstein (2006) and Blonigen and Soderbery (2010) are quite different from 1.

[23] Values of σ from Broda and Weinstein (2006, Table V) vary from 1.2 for footwear (very differentiated) through 17.1 for crude oil (very homogenous). The corresponding ρ values are 0.17 and 0.94. Coffee comes in at $\rho = 0.6$: the CES would then suggest that unit costs for coffee should be close to uniform if (7) is to hold and the revenue tail parameter is close to 1 in the industry.

$\eta_{f_V}(v) > -1$ *iff* $\theta < 0$. *Furthermore,* $\eta_{f_V}(v)$ *is strictly decreasing in* θ *iff* $\eta_{f_U}(v) > -1$.

Proof. Consider first $\theta > 0$. Because $F_V(v) = F_{kU^\theta}(v) = F_U\left[\left(\frac{v}{k}\right)^{1/\theta}\right]$, then we have

$$f_V(v) = \frac{1}{\theta}\left(\frac{v}{k}\right)^{1/\theta}\frac{1}{v}f_U\left[\left(\frac{v}{k}\right)^{1/\theta}\right].$$

Differentiating this identity,

$$\begin{aligned} f_V'(v) &= \frac{1}{\theta^2}\left(\frac{v}{k}\right)^{2/\theta}\frac{1}{v^2}f_U'\left[\left(\frac{v}{k}\right)^{1/\theta}\right] + \left(\frac{1}{\theta}-1\right)\frac{1}{\theta}\left(\frac{v}{k}\right)^{1/\theta}\frac{1}{v^2}f_U\left[\left(\frac{v}{k}\right)^{1/\theta}\right], \\ \text{so} \qquad \eta_{f_V}(v) &= \frac{1}{\theta}\left\{\left(\frac{v}{k}\right)^{1/\theta}\frac{f_U'\left[\left(\frac{v}{k}\right)^{1/\theta}\right]}{f_U\left[\left(\frac{v}{k}\right)^{1/\theta}\right]} + (1-\theta)\right\}. \end{aligned}$$

Hence the required condition (14) holds.

Similarly, for the negative power transformation (i.e., when $\theta < 0$) in Lemma 1*(c)*, because $F_V(v) = F_{kU^\theta}(v) = 1 - F_U\left[\left(\frac{v}{k}\right)^{1/\theta}\right]$, we again obtain (14). Both cases ($\theta > 0$ and $\theta < 0$) are covered by the statement in the Proposition.

For the last part, assume first that $\eta_{f_U}(v) > -1$. In this case, the relation $\eta_{f_V}(v) = \frac{1}{\theta}\left\{\eta_{f_U}(v) + 1\right\} - 1$ defines two decreasing rectangular hyperbolae separated by $\theta = 0$ and $\eta_{f_V}(v) = -1$. If $\eta_{f_U}(v) < -1$, the two rectangular hyperbolae are increasing, with the same asymptotes. ■

Notice from (14) that $\eta_{f_U}(v) = -1$ if and only if $\eta_{f_V}(v) = -1$, regardless of $\theta \neq 0$. Furthermore, $\eta_{f_V}(v)$ increases with $\eta_{f_U}(v)$ if and only if $\theta > 0$, so that density elasticities move in the same direction when the economic variables are positively related.

The uniform distribution for U clarifies the role of the other term. Then $\eta_{f_U} = 0$, and so $\eta_{f_V}(v) > 0$: hence, $f_V'(v) > 0$ if and only if $\theta \in (0,1)$. Then an increasing $f_U(u)$ is reinforced,

but a decreasing one is offset (and may be overturned). That is, $f_V'(v) \geq 0$ is guaranteed for $f_U'(u) \geq 0$ and $\theta \in (0,1)$; and $\theta > 1$ is necessary for $f_V'(v) \leq 0$ if $f_U'(u) \geq 0$. For $\theta < 0$, $f_V'(v) < 0$ if $f_U'(u) \geq 0$. For example, uniform costs in CES imply a decreasing revenue density (due to convexity of the revenue function in c).

More specifically, taking U as cost, then elasticities of output (where $\theta = -\frac{1}{1-\rho} < 0$) and revenue are (where $\theta = -\frac{\rho}{1-\rho} < 0$):[24]

$$\eta_{f_Y} = -(1-\rho)\left(\eta_{f_C} + 1 + \frac{1}{1-\rho}\right) \tag{15}$$

and

$$\eta_{f_R} = -\left(\frac{1-\rho}{\rho}\right)\left(\eta_{f_C} + \frac{1}{1-\rho}\right). \tag{16}$$

For the Pareto circle, the output and revenue density elasticities are $\eta_{f_Y} = -(\alpha_y + 1)$, and $\eta_{f_R} = -(\alpha_r + 1)$ respectively.[25] The cost (or price) density elasticity is $\eta_{f_C} = \alpha_{\hat{c}} - 1$ (so that η_{f_C}must exceed -1 for the underlying cost distribution to be increasing in c). Recall that the cost density elasticity is the same as the price one, so cost and price can be interchanged in the statements below. We prefer to retain costs because they are a primitive to the model.

Notice first that a greater cost density elasticity feeds through to a smaller output or revenue density elasticity. Recalling that the latter are negative, this means that those densities become more elastic (more responsive to output or profit levels): more variation in the fundamental variable causes more variation in the induced economic variables.

[24] Or, in terms of the elasticity of substitution, $\sigma\eta_{f_Y} = -\eta_{f_C} - 1 - \sigma$ and $\eta_{f_\Pi} = -\frac{1}{\sigma-1}\left(\eta_{f_C} + \sigma\right)$.

[25] The output and profit ones are elastic (below -1) because the α's are positive. The cost one must exceed -1 (but can be positive) because $\alpha_{\hat{c}} > 0$ for the cost distribution to be increasing.

The impact on these elasticities of the degree of industry product differentiation is rather interesting. The derivatives of (15) and (16) with respect to ρ have the sign of $\left(\eta_{f_C}+1\right)$, which is positive. This means that the density elasticities are larger, and (being negative) hence actually *less* elastic when there is less product differentiation (higher ρ). This is because the equilibrium profit and output functions become more elastic (as expected) when ρ rises, since the underlying cost heterogeneity is parlayed into more revenue variability when products in the industry are closer substitutes.[26] Then revenue and output densities (i.e., as functions of revenue and output respectively) get less responsive to their arguments. These densities are flatter because there is higher variability in the outputs and revenues of firms.[27]

Combining (15) and (16) gives the relation between output and revenue density elasticities as

$$\eta_{f_R}=\frac{1}{\rho}\eta_{f_Y}+\frac{1-\rho}{\rho}. \tag{17}$$

This indicates that these two elasticities are positively related, with the revenue one both more leveraged (i.e., its coefficient, $\frac{1}{\rho}>1$) and more responsive to the output one the smaller is ρ. Moreover, recalling that both elasticities are below -1,[28] a higher ρ entails a higher (less elastic) revenue density elasticity: less product differentiation in the sector means a flatter revenue density even conditioning on a given output density.

[26] For example, the profit elasticity is $\frac{\rho}{\rho-1}$, so its derivative is negative. The elasticity is negative, and becomes more so as ρ increases.

[27] This can be seen most clearly with a uniform cost density, so $\eta_{f_C}=0$. Then the density elasticities are $\eta_{f_Y}=-\left(2-\rho\right)$ and $\eta_{f_r}=-\frac{1}{\rho}$ and so higher ρ increases both and they become less responsive. If ρ is small, the CES demand system entails that revenues and outputs are almost the same for all firms, so densities pile up on very close output or revenue values.

[28] The derivative with respect to ρ has the sign of $-\left(\eta_{f_Y}+1\right)$, which is positive because $\eta_{f_Y}<-1$.

4.3 Modes and slopes

Equations (15) and (16) imply several slope conditions, including:

$$f_Y'(y) < 0 \text{ iff } \eta_{f_C}(y) > -1 - \frac{1}{1-\rho} \text{ and } f_R'(r) < 0 \text{ iff } \eta_{f_C}(r) > -\frac{1}{1-\rho}.$$ [29]

These relations combine technological and taste distribution properties. A decreasing revenue density implies a decreasing output density, while an increasing output density is necessary for an increasing revenue density (since $f_R'(r) < 0$ iff $\eta_{f_Y} < \rho - 1$ by (17)). The relations also have implications for properties of distribution modes. For example, if both output and revenue densities were uni-modal, then the modal revenue level is higher than the revenue level associated to the modal output. Loosely, the most common revenue level is higher than the revenue of the most common output level: if this were not true in the data, the data could not have been given from a CES model. Also, note from (15) and (16) that an increasing cost density implies both output and revenue densities are decreasing. Recall though that high costs are associated to low output and revenue, so if the cost density is uni-modal then the others can also be uni-modal.[30]

Likewise, an analogous reasoning with (17) indicates that if the output density has a single peak then the revenue density function is increasing at the corresponding revenue, so the mode of the revenue density is below the revenue level of the modal density. Pulling this together, the output density peaks before the revenue density, which in turn peaks before the cost density.

[29] Here the η_{f_C}'s depend on c through y and π respectively. When we compare density shapes below, we are comparing at the y and π values that are compatible through the same c that generates them from (3) and (5).

[30] For the power cost case, the associated revenue and output densities are always decreasing (as should be the case since they are Pareto!) To see this, note that for $F_C(c) = (c/\underline{c})^{\alpha_{\hat{c}}}$ then $\eta_{f_c} = \alpha_{\hat{c}} - 1$ and so the revenue slope condition above becomes $f_R'(r) < 0$ iff $\alpha_{\hat{c}} + \frac{\rho}{1-\rho} > 0$, which must hold since both terms are positive.

Put another way, decreasing cost density drives decreasing revenue density drives decreasing output density. The chain reverses for increasing output density driving the two others increasing. These qualitative features should hold empirically under the CES when there is enough information on two or three densities.

Proposition 2 above implies that the standard CES model with cost heterogeneity alone cannot deliver (say) Pareto distributions for *both* revenue and prices. Indeed, if revenue is Pareto distributed, then price must follow a power distribution. We next introduce quality heterogeneity in Section 5 to decouple the CES-circle into two or more satellite orbits, and so enrich the associated distribution circles.

5 CES quality-enhanced model

5.1 General specification

We now extend the model to allow for quality differences across products. Quality heterogeneity appears to play an important role empirically. Following Baldwin and Harrigan (2011) and Feenstra and Romalis (2014), we rewrite the Representative Consumer's sub-utility functional as $\chi = \left(\int_{\Omega} z(\omega)^{\rho} d\omega\right)^{1/\rho}$ with $\rho \in (0,1)$ and interpret $z(\omega) = v(\omega) q(\omega)$ as the *quality-adjusted consumption* of variant ω, where $v(\omega)$ is its quality and $q(\omega)$ is the quantity consumed (as before). We clarify the quality interpretation in Section 6 in terms of a disaggregate model of individual discrete choice. The corresponding demands are:

$$h(p_i, \hat{p}_i) = \frac{\Xi(I)}{p_i} \frac{\hat{p}_i^{\frac{\rho}{\rho-1}}}{\int_{\omega \in \Omega} \hat{p}(\omega)^{\frac{\rho}{\rho-1}} d\omega}, \tag{18}$$

where we have defined $\hat{p}_i = p_i/v_i$, which is interpreted as the price per unit of "quality," and $\Xi(I)$ is as in Section 2 for the three different cases (the amount spent on the differentiated commodity).

The key feature of (18) is that p_i enters both with and without quality. The standard model (1) ensues when all the v's are the same.

With a continuum of firms (as per the usual monopolistic competition set-up), Firm i's equilibrium price solves $\max_{p_i} \left\{ \frac{(p_i - c_i)}{p_i} \hat{p}_i^{\frac{\rho}{\rho-1}} \right\}$, so that the pricing solution $p_i = \frac{c_i}{\rho}$ and the Lerner condition (2) still hold independent of quality. Hence, defining $x_i = v_i/c_i$, which we refer to as quality/cost, all firms set the same proportional mark-up, and the equilibrium profit now depends on quality/cost:

$$\pi_i = (1-\rho)\,\Xi(I)\,\frac{x_i^{\frac{\rho}{1-\rho}}}{D_X} = (1-\rho)\,r_i, \tag{19}$$

where $D_X = \int_{\omega\in\Omega} x(\omega)^{\frac{\rho}{1-\rho}}\,d\omega$ is assumed to converge. Equilibrium profit is still a fraction $(1-\rho)$ of sales revenue, so their two distributions are the same up to a scale factor. Likewise, (19) implies that profit, sales revenue, and quality/cost distributions are in the same closed power-family.[31]

Price and cost distributions are still the same up to a scale factor, but productivity (reciprocal costs) and profits are no longer necessarily in the same power-family (because one depends on the distribution of c and the other on the distribution of x). How the cost and profit distributions are linked is determined by the link between cost and quality. A functional relation between cost and quality/cost ties this down (and is illustrated in the next sub-section), along

[31] Profits are increasing in x so that firms would like this as large as possible. We can link cost and quality through a type of production function and have (heterogeneous) firms choose their x. More anon.

with the other distributions on the profit side.

The equilibrium output is given by

$$y_i = \frac{\rho}{c_i} \Xi (I) \frac{x_i^{\frac{\rho}{1-\rho}}}{D_X}, \tag{20}$$

and so the output distribution may be in neither the productivity nor the quality/cost power-family, because it draws from both the cost and quality/cost distributions (none of these links are explored in Baldwin and Harrigan, 2011, or Feenstra and Romalis, 2014.)

Therefore, there are at most three (linked) distribution families. To summarize:

Proposition 4 *(Decoupling the CES circle) Consider the quality-enhanced CES model of monopolistic competition. Then, in equilibrium:*

(a) Price and cost distributions are multiplicatively related;

(b) Profit and sales revenue distributions are multiplicatively related;

(c) Profit, sales revenue, and quality/cost distributions are positive-power related;

(d) The output distribution is not in general multiplicatively or power related to any of the other distributions.

We can derive analogous elasticity relations to those in Section 3 to link the elasticities of the densities on the third leg (c).[32] We next illustrate the Proposition with an example which specifies a power relation between quality and cost and amends statement (d) of the Proposition.

[32] Namely, $\eta_{f_Y} = (1-\rho)\left(\eta_{f_X} - 1 - \frac{1}{1-\rho}\right)$ and $\eta_{f_\Pi} = \left(\frac{1-\rho}{\rho}\right)\left(\eta_{f_X} - \frac{1}{1-\rho}\right)$.

5.2 Constant elasticity quality/cost relation

Suppose that $x = c^{\gamma}$ so that quality/cost is increasing with cost if $\gamma > 0$ (i.e., quality rises faster than cost), and it is decreasing if $\gamma < 0$. The latter case is embodied in the standard CES model above where $\gamma = -1$ and so "better" firms are those with lower costs. The former case effectively corresponds to Feenstra and Romalis (2014).[33] The advantage of the constant elasticity relation is that it allows us to deploy results (Lemma 1) on applying power transforms to random variables.

Because revenues are proportional to $x_i^{\frac{\rho}{1-\rho}}$ (see (19)), they are proportional to $c_i^{\gamma\frac{\rho}{1-\rho}}$. Hence if $\gamma > 0$, the revenue distribution is in the same power-family as the cost distribution. So then too are the profit and quality-cost distributions (see (3)). But if $\gamma < 0$, profits, revenues and quality-costs are in the inverse (or "opposite") power-family, which is the power-family of productivity.[34] This is the generalization of the earlier standard CES result. Prices, of course, still distributed (up to a scale factor) like costs.

Output is more intricate because it draws its influences from both sides. Indeed, from (18), output is proportional to $x_i^{\frac{\rho}{1-\rho}}/c_i$, which equals $c_i^{\gamma\frac{\rho}{1-\rho}-1}$ under the constant elasticity quality/cost formulation. This implies that for $\gamma > \left(\frac{1-\rho}{\rho}\right)$ the output and cost distributions are in the same power-family, while otherwise they are in inverse families. A summarizing statement:

Proposition 5 *(Constant elasticity quality/cost relation) Consider the quality-enhanced CES*

[33] Along the same lines as Feenstra and Romalis (2014), we can let $v = l^{\alpha}$ be the quality produced at cost $wl+\phi$ with ϕ a firm-specific productivity shock, where l is labor input, w is the wage, and $\alpha \in (0,1)$. Maximizing $x = l^{\alpha}/(wl+\phi)$ gives the optimized value relation between cost and quality as $x = \left(\frac{\alpha}{w}\right)^{\alpha} c^{\alpha-1}$ and so the quality realtion takes a power form. Here it is decreasing (and depends on the fundamental via $x = \phi^{\alpha-1}$).

[34] Recall that one distribution is the *inverse* of another one if it is its survival function.

model of monopolistic competition with $x = c^{\gamma}$. Then, in equilibrium:

(a) Price and cost distributions are multiplicatively related;

(b) Profits, sales revenue, and quality/cost distributions are positive-power related to unit costs for $\gamma > 0$ and negative-power related for $\gamma < 0$;

(c) The output distribution is negative-power related to unit costs for $\gamma < \left(\frac{1-\rho}{\rho}\right)$, and positive-power related for $\gamma > \left(\frac{1-\rho}{\rho}\right)$.

The actionable tests for the model are that estimated profit and revenue distributions should be the same up to a multiplicative factor. Likewise for prices and cost.[35] If these hold (which is also true for Proposition 4), the test for the constant elasticity bridge function is that these distributions lie in the same CPF. If so, then the value of γ can be estimated from them. And the output distribution should be in the same CPF too.

Take the example of a Pareto distribution for unit costs. First, prices are also Pareto distributed. Second, profits, revenue, and quality/cost are Pareto distributed for $\gamma > 0$ and power distributed for $\gamma < 0$ (they are independent of cost if $\gamma = 0$). Third, output is power distributed for $\gamma < \left(\frac{1-\rho}{\rho}\right)$, and Pareto distributed for $\gamma > \left(\frac{1-\rho}{\rho}\right)$.[36] Hence, *this formulation can deliver Pareto distributions for both prices and profits via the constant elasticity quality/cost function.* Pareto revenue and profit distributions are well documented in the literature. Coad (2009) analyzes price distributions for wine, one-week holidays in Majorca, used cars, and London house prices, situations where there is considerable quality differentiation.[37] He finds

[35] This is equivalent to mark-ups being constant for the CES: this restriction we soften in our demand extension Section 6.

[36] If costs are power distributed, Pareto and power are reversed in the above statements.

[37] Note though that for the first two cases the products are sold by a single firm.

the resulting distribution is close to Pareto, though less skewed (and more skewed than the lognormal).

Proposition 5(b) indicates that quality/cost and profit distributions are in the same power-family. For example, suppose that the distribution of quality/costs is Pareto: $F_X(x) = 1 - (\underline{x}/x)^{\alpha_x}$ and assume that $\alpha_x \frac{1-\rho}{\rho} > 1$, so that the size distribution of profit is Pareto with tail parameter $\alpha_\pi = \alpha_x \frac{1-\rho}{\rho}$. Our result is that the profit tail parameter is the product of a preference parameter and a quality/cost distribution one.[38]

While the introduction of quality enables matching of disparate distribution pairs, the CES still involves at most three distribution families (only two for the constant elasticity case, with one being the inverse family of the other). Second, prices are independent of qualities, but percentage cost increases are passed on at equal percentage rates because the Lerner measure, $l = \frac{p-c}{p}$, is constant. Even when c depends on quality, l is independent of quality.

Atkin et al. (2015) find that the elasticity of mark-ups is significantly greater than that for costs, and quality differences appear to play an important role in this. Therefore we next introduce a generalization of the structure above, which might better explain the patterns found, and it yields flexible mark-up patterns.

6 (Really) Beyond the CES

So far, we have considered distributions of equilibrium variables on the supply side. These are driven by fundamental productivity (and/or quality) heterogeneity. The CES model is usually presented from a Representative Consumer utility perspective. However, individuals

[38] Although why they yield the same constant across settings remains intriguing.

typically consume one (or at most a few) of product variants. The question whether the representative consumer can capture an aggregate relation of individuals making discrete choices was addressed by Anderson, de Palma, and Thisse (1992) who showed how to underpin the CES demand system by aggregating heterogeneous individuals' discrete choices. Here, we micro-found how quality enters individual utilities consistent with aggregating to the previous formulation, and use this technique to generalize the demand model to bridge CES to Logit, which is the workhorse of empirical IO. We construe cases in between and beyond both book-end cases. Doing so enables us to deliver a wide variety of mark-up properties: empirical work on mark-ups can identify the appropriate demand model.

6.1 Discrete choice roots for CES and Logit

There is a set of consumers (with mass 1) making choices according to a discrete choice framework. Suppose that each consumer makes a discrete choice of which variant to buy, based on a conditional (indirect) utility for variant i as

$$u_i = \bar{u}_i + \varepsilon_i, \qquad i \in \Omega \tag{21}$$

where $\bar{u}_i$ is the common "measured" utility and ε_i is the individual preference shock (or idiosyncratic draw). We assume that $\bar{u}_i$ has a Constant Relative Risk Aversion (Box-Cox transformation) form, so:

$$\bar{u}_i = \bar{I} + v_i - \frac{p_i^b - 1}{b}, \qquad i \in \Omega, \tag{22}$$

where $\bar{I}$ is individual income, v_i represents the quality of option i, $\bar{u}_i$ is decreasing in p_i, and b is a parameter enabling us to move seamlessly between CES ($b = 0$) and Logit ($b = 1$). Each consumer chooses the option for which u_i is greatest. Assume that the ε_i are i.i.d. Type 1 Extreme Value (or "Gumbel") distributed with scale parameter $\mu \geq 0$.[39] The choice probabilities for any *option* follow the continuous version of the standard logit formula:

$$\mathbb{P}_i = \frac{\exp\left(\frac{\bar{u}_i}{\mu}\right)}{\int_{\omega\in\Omega} \exp\left(\frac{\bar{u}(\omega)}{\mu}\right) d\omega}. \tag{23}$$

Applying Roy's identity to (22) indicates that the quantity consumed of option i, conditional on preferring it, is p_i^{b-1}. Therefore, the expected demand for i is $h(p_i) = p_i^{b-1}\mathbb{P}_i$ or

$$h(p_i) = p_i^{b-1} \frac{\exp\left(\frac{\bar{u}_i}{\mu}\right)}{\int_{\omega\in\Omega} \exp\left(\frac{\bar{u}(\omega)}{\mu}\right) d\omega}. \tag{24}$$

6.2 Case $b = 0$: CES

For $b = 0$ when all v_i are the same, we obtain the CES demand form (1) from l'Hôpital's rule:

$$h(p_i) = \frac{1}{p_i} \frac{p_i^{-\frac{1}{\mu}}}{\int_{\omega\in\Omega} p(\omega)^{-\frac{1}{\mu}} d\omega}, \tag{25}$$

where the parameters are matched by $\mu = \frac{1-\rho}{\rho} = \frac{1}{\sigma-1}$ (or $\rho = \frac{1}{1+\mu}$). Notice how the limit cases concur. If $\mu \to 0$, products are perfect substitutes from (21) and this corresponds to $\rho \to 1$ for the CES. If $\mu \to \infty$, then idiosyncratic tastes are paramount, and $\rho \to 0$ (the Cobb-Douglas limit).

[39] The Gumbel distribution takes the form $F(s;\beta,\mu) = \exp(-\exp(\beta - s)/\mu)$, where β is the "location" parameter. The mean is $\beta + \mu\gamma$ (where γ is Euler's constant), and the standard deviation is $\pi\mu/\sqrt{6}$.

For $b = 0$ and different v_i, and letting $x_i = \tilde{v}_i / p_i$, where $\tilde{v}_i = \exp v_i$ is a convenient rescaling of measuring quality, we recover the form of (20):

$$h(p_i) = \frac{1}{p_i} \frac{x_i^{\frac{1}{\mu}}}{\int_{\omega \in \Omega} x(\omega)^{\frac{1}{\mu}} d\omega}, \tag{26}$$

with $\mu = \frac{1-\rho}{\rho}$.

We now describe the relation between equilibrium output and quality that comes from the CES model. Temporarily dropping subscripts to ease clutter, we have output and price (recalling the earlier CES mark-up formula, (2)) respectively as

$$y = \frac{1}{D_X} \frac{1}{p} \left(\frac{\tilde{v}}{p} \right)^{1/\mu} \text{ with } p = (1+\mu)\, c(\tilde{v}), \text{ so } y = \frac{1}{D_X} \tilde{v}^{1/\mu} \left[(1+\mu)\, c(\tilde{v}) \right]^{-(1+\mu)/\mu}, \tag{27}$$

where we make explicit the dependence of unit cost on quality, and where the equilibrium value of the denominator in (26) is $D_X = (1+\mu)^{-1/\mu} \int_{\omega \in \Omega} \left[\tilde{v}(\omega) / c(\omega) \right]^{\frac{1}{\mu}} d\omega$. Then $l = \frac{p-c}{p} = \frac{\mu}{1+\mu}$ is independent of c.

Differentiation yields $\frac{dy}{d\tilde{v}} \propto \left[c(\tilde{v}) - \tilde{v} c'(\tilde{v}) \right]$, so that output rises if the cost function is inelastic. This result stems from two conflicting forces. Higher quality raises demand at constant prices, but raises price too through the mark-up. If cost rises quickly enough with quality, then output falls.

These forces are present in the extended model too, so that whether or not higher size is associated to higher quality depends on the elasticity of the cost function. Atkin et al. (2015) find (for the case of footballs made in Sialkot, Pakistan) that mark-ups rise with firm size, as measured by employment. Since they argue that higher quality is expressed as higher quality inputs, then we can take employment as a good proxy for output. Thus we could infer, through

the lens of the current model, that costs are quite inelastic in quality.

6.3 Case $b = 1$: Logit

When $b = 1$ we have the standard (quality-enhanced) Logit model with[40]

$$h(p_i) = \frac{\exp(x_i/\mu)}{\int_{\omega\in\Omega} \exp(x(\omega)/\mu)\, d\omega}, \tag{28}$$

and here $x_i = v_i - p_i$. Straightforward algebra yields the equilibrium price for the monopolistically competitive Logit model is

$$p = c(v) + \mu, \tag{29}$$

which concurs with the limit of the oligopoly model (given in Anderson et al., 1992) when the number of firms gets large. We have thus just introduced the tractable and simple Logit monopolistic competition model into the literature.

6.4 Case $b \in (0, 1)$: between Logit and CES

We now describe the equilibrium properties for $b \in (0, 1)$.[41]

Firm i's profit is proportional to

$$\pi_i = (p_i - c_i)\, p_i^{b-1} \exp\left(\frac{\bar{u}_i}{\mu}\right),$$

where we recall from (22) that $\bar{u}_i = \bar{I} + v_i - \left(p_i^b - 1\right)/b$. Then we have

$$\frac{d\pi_i}{dp_i} = p_i^{b-1}\left(\frac{p_i^b}{\mu} - b + 1\right) \exp\left(\frac{\bar{u}_i}{\mu}\right) \left\{\frac{1}{\frac{p_i^b}{\mu} - b + 1} - l_i\right\}, \tag{30}$$

[40] An outside option is readily appended here and below.
[41] We discuss below what happens outside the range $b \in (0, 1)$.

where the terms before the parentheses are positive (for $p_i > 0$) and recall $l_i = (p_i - c_i)/p_i$ is the Lerner index. Here l_i monotonically increases from 0 at $p_i = c_i$ to 1 when $p_i \to \infty$. The term $1/\left(\frac{p_i^b}{\mu} - b + 1\right) > 0$ is strictly decreasing in $p_i > 0$ and asymptotes to 0 as $p_i \to \infty$. Thus there is a unique solution to the first-order condition $\frac{d\pi_i}{dp_i} = 0$, call it p_i^e. Moreover profit is strictly quasi-concave, because $\frac{d\pi_i}{dp_i} > 0$ when $p_i < p_i^e$ and $\frac{d\pi_i}{dp_i} > 0$ when $p_i > p_i^e$. When $b \to 1$ (Logit), the equilibrium price tends to (29), while the CES price form in (27) results when $b \to 0$.

The comparative static effects of an increase in c_i are to decrease equilibrium l_i, as can be seen by noting that l_i moves down for given p_i, while the other term in brackets in (30) is unchanged. However, a change in quality, v_i, does not alter the equilibrium percentage mark-up, l_i. The implication is that when higher quality is associated to higher cost, then percentage mark-up falls. We discuss below some empirical evidence for a particular industry, and how this can be consistent with $b < 0$ instead.[42]

The absolute mark-up, $m_i = p_i - c_i$, moves the other direction with c_i (for $b \in [0, 1)$). To see this, rewrite the profit derivative as

$$\frac{d\pi_i}{dp_i} = p_i^{b-2}\left(\frac{p_i^b}{\mu} - b + 1\right)\exp\left(\frac{\bar{u}_i}{\mu}\right)\left\{\frac{p_i}{\frac{p_i^b}{\mu} - b + 1} - m_i\right\} \tag{31}$$

and note that that the slope of the first term in brackets has the sign of $(1 - b)$ and that this slope is below 1. Again, quality only impacts the mark-up through the cost, and the positive slope property implies that m_i rises with c_i.

[42] Different parameter values could hold for different industries, of course, and our model ranges over a wide span.

6.5 Case $b > 1$: the super-logit

While this case is incongruent with the discrete choice underpinnings given above (because the conditional demand slopes up), it nonetheless delivers a system of product *demands* that slope down under some parameter restrictions, and allows us to analyze a tractable functional form for demand that goes the "other side" of Logit, where absolute mark-ups decrease with cost.

The product demand slopes down for $p_i > c_i$ as long as $c_i > ((b-1)\mu)^{1/b}$, and so we henceforth assume that condition holds. Now, along the lines of the earlier analysis, we have that profit remains quasi-concave and that higher c_i concurs with a lower percentage mark-up, l_i, because the first term in brackets in (31) is decreasing for $b > 1$. However, now it means a lower absolute mark-up, m_i, too. Clearly, the latter property implies the former.[43]

6.6 Case $b < 0$: the sub-CES

Because the first term in parentheses in (30) is flat for the CES, then profit quasi-concavity still holds in the neighborhood of $b = 0$. Surprisingly, this property is preserved for all $b < 0$. To see this, recall that profit is quasi-concave if the demand reciprocal is convex (Caplin and Nalebuff, 1991). The demand reciprocal is proportional to $p^{1-b}\exp\left(\frac{p^b}{\mu b}\right)$, which has second derivative proportional to $-b(1-b)p^{-b-1} + \frac{(1-b)}{\mu}p^{-1} + \frac{1}{\mu^2}p^b$; each term is positive for $b < 0$, as desired. Now higher c_i necessarily entails higher percentage mark-up, l_i, for now the term $\frac{1}{\frac{p_i^b}{\mu}-b+1}$ in (30) increases in c_i. Higher l_i drives higher m_i too.

The implication here is that higher qualities, through higher production costs, now drive higher mark-ups. The regime $b < 0$ is consistent with Finding (3) of Atkin et al. (2015)

[43] A lower m_i implies a lower l_i, and the contrapositive is that a higher l_i implies a higher m_i: see Table 1 below.

insofar as they argue that mark-up elasticity exceeds cost elasticity with respect to firm size (as measured by employment). Their Finding 4 is that larger firms have higher costs due to higher quality inputs, so higher mark-ups are set on high quality products (Finding 5). In order to generate the finding that large size correlates with high quality we refer back to the property enunciated in Section 6.2 that higher quality raises demand but raises price too, but the direct effect dominates in equilibrium when cost rises slowly enough with quality. So our takeaway is that these findings are consistent with $b < 0$ and there being an inelastic relation between costs and quality.

We conclude with a summary Table for the two different mark-ups.

	$b < 0$ sub-CES	$b = 0$ CES	$b \in (0, 1)$	$b = 1$ Logit	$b > 1$ super-Logit
l	$\uparrow$	$\frac{\mu}{1+\mu}$	$\downarrow$	$\frac{\mu}{c+\mu}$	$\downarrow$
m	$\uparrow$	$c\mu$	$\uparrow$	μ	$\downarrow$

When relations between economic variables are monotone, the analysis goes through as before. In the discussion below, we assume that quality is fixed, as in the CES standard model and as in Mrázová et al. (2017). Equilibrium profit is decreasing in cost (by the envelope theorem). The Lerner index and markup are monotone transformations, increasing or decreasing according to the value of b (see Table). Output decreases with price (a condition on marginal cost is required when $b > 1$). Finally, equilibrium price increases with cost (see (31)). These monotonicity properties translate into one-to-one relations between CDFs (or survival functions) of cost, output, price, mark-up, and profit, using similar relations as before. Here the CDF of price is numerical, since the equilibrium price does not have a closed form (see (31))

except for CES and Logit, which have explicit prices. We next treat the Logit with quality and provide explicit relations between the CDFs.[44]

7 Distributions for Logit

We finish up with the analysis of distributions for the Logit, which corresponds to $b = 1$. The Logit model is a mainstay for structural empirical Industrial Organization. The monopolistic competition model delivers closed-form solutions for the primary variables.

The key features of Logit are that absolute mark-up is constant, $p_i = c_i + \mu$ (from (29)) so that higher qualities bear the same mark-up as lower ones, although firm size is consequently larger. The equilibrium price distribution therefore tracks the unit cost distribution (up to the additive constant, μ). Because then $\pi_i = m_i y_i = \mu y_i$, the profit and output distributions are in the same multiplicative class. The revenue distribution depends on both the cost distribution and the quality-cost distribution (recall $x_i = v_i - c_i$). The most interesting relation is between the profit (or output) and quality-cost distributions.

The following extension to Lemma 1 gives the relevant new transformations.

Lemma 3 *(Transformation extension) Let $F_U(u)$ be the CDF of a random variable U. Then, the CDF of $F_V(v)$ for the additive transformation ($V = k(U + a)$, $k > 0, a > 0$), exponential transformation ($V = k_1 \exp(k_2 U)$, $k_1 > 0, k_2 > 0$) are respectively given by:*

(a) (Additive) $F_V(v) = F_U\left(\frac{v}{k} - a\right)$;

[44] Anderson and de Palma (2001) rank firms by quality-cost for the Logit oligopoly. The concept needs to be extended for the generalized CES model to determine the monotonicity relations needed to derive relations between the CDF of the different economic variables (see the interpretation of quality by Kugler and Verhoogen, 2012, Appendix D).

(b) (Exponential) $F_V(v) = F_U\left[\frac{1}{k_2}\ln\left(\frac{v}{k_1}\right)\right]$.

The importance of the Lemma is that both profit and output are related to quality-cost by exponential transforms for the Logit. This transformation renders log distributions. First, the *normal* distribution is perhaps the most natural primitive assumption to take for quality-costs. Then profit $\Pi \in (0, \infty)$ is *log-normally* distributed. The log-normal has sometimes been fitted to firm size distribution (see Cabral and Mata, 2003, for a well-cited study of Portuguese firms). Note that a truncated normal begets a truncated log-normal.

The simplest text-book case is the *uniform* distribution. Then, by Lemma 3b the equilibrium profit Π has distribution $F_\Pi(\pi) = \mu \ln\left(\frac{\pi D_L}{\mu}\right)$, where D_L is the logit denominator, and the profit density is unit elastic. A *truncated Pareto* distribution leads to a truncated Log-Pareto for profit (or output).

The most successful function to fit the distribution of firm size has been the Pareto. The corresponding distribution of quality-cost ($x = v - c$) is:

Proposition 6 *(Exponential and Pareto distributions under Logit) For the Logit model of monopolistic competition, and let quality-cost be exponentially distributed:* $F_X(x) = 1 - \exp[-\lambda(x - \underline{x})]$, $\lambda > 0$, $\underline{x} > 0$, $x \in [\underline{x}, \infty)$, *with* $\lambda\mu > 1$. *Then equilibrium output and profit are Pareto distributed:* $F_Y(y) = 1 - \left(\frac{\underline{y}}{y}\right)^{\alpha_y}$ *and* $F_\Pi(\pi) = 1 - \left(\frac{\underline{\pi}}{\pi}\right)^{\alpha_\pi}$, *where* $\alpha_y = \alpha_\pi = \lambda\mu > 1$.

The proof is in the Appendix. Thus the shape parameter, $\alpha_y = \alpha_\pi$, for the endogenous economic distributions depends just on the product of the taste heterogeneity and the technology shape parameter. Note that D_L is bounded if $\mu > 1/\lambda$, i.e., the taste heterogeneity should exceed average quality-cost.

8 Conclusions

The CES model has been the workhorse model of monopolistic competition with asymmetric firms. The central distribution in the literature has been the Pareto for sales revenues. We show that all relevant distributions are Pareto if any one is (caveat: for prices and costs it is the distribution of the reciprocal that is Pareto). This result we term the *Pareto circle*. To put this another way, suppose we posit that productivity (the reciprocal of costs) is Pareto distributed. (Equivalently, costs have a power distribution.) Then, so does the reciprocal of prices follow a Pareto distribution, and the other variables (output, revenue, and profit) are all Pareto distributed. The Pareto circle cannot be escaped if one element is Pareto. It is not possible to have (for example) a Pareto distribution for profits and (another) Pareto distribution for prices in the CES model. Similar results hold for other distributions, yielding a more general *CES circle*: the assumed distribution of productivity (cost reciprocal) is in the same closed power-family as the equilibrium distributions of outputs, profits, etc.

This analysis determines the logical relations between equilibrium distributions. These are simple for the core Pareto distribution, but they turn out to be surprisingly general once couched as relations between elasticities of equilibrium densities. Put another way, the Pareto results form a solid benchmark for broader distribution relations. The relations (between equilibrium distributions) describe how shapes of densities, as described by their elasticities, are all related to each other in the CES through a simple linear relation.

The CES circle is broken by allowing for a further dimension of firm heterogeneity. Following Baldwin and Harrigan (2011) and Feenstra and Romalis (2014), we introduce "quality"

heterogeneity and link the quality and cost distributions via a function that writes quality as a function of cost. Doing this then enables us to get linked *groups* of distributions. In one group are profit, revenue, and quality/cost; and in another are costs and prices. The output distribution draws from both these groups. Our leading example is a quality/cost function that can deliver Pareto distribution for both revenues *and* prices.

Even with the introduction of quality, the CES remains restrictive in the properties it imposes. To break the constant (percentage) mark-up property we introduced a simple parameterization linking CES to Logit and forms beyond. One virtue of the extended model is that it helps explain recent results on quality, firm size, and mark-ups. With more data coming onstream, more empirical verification of the shapes of the distributions of these variables should be forthcoming. Such data would evaluate the scope of the proposed extended CES, and suggest whether more elaborate demand models are required to better explain observed economic distributions.

9 Appendix

Proof of Proposition 6

We first calculate the logit denominator, D_L, from (28), using the exponential CDF $F_X(x) = 1 - \exp(-\lambda(x - \underline{x}))$ with density $f_X(x) = \lambda \exp(-\lambda(x - \underline{x}))$ and $\lambda > 0$, $\underline{x} > 0$, and $x \in [\underline{x}, \infty)$. Integrating,

$$D_L = \frac{M\lambda\mu}{\lambda\mu - 1} \exp\left(\frac{\underline{x}}{\mu}\right),$$

which is positive and bounded under the assumption that $\lambda\mu > 1$. Now, the CDF of Π is given

by $F_{\Pi}(\pi) = \Pr\left(\frac{\exp(x/\mu)}{D_L} < \pi\right) = \Pr\left(x < \mu \ln\left(D_L \pi\right)\right) = F_X\left(\mu \ln\left(D_L \pi\right)\right)$. Then

$$F_{\Pi}(\pi) = 1 - \exp\left(-\lambda\left(\mu \ln\left(\frac{\pi D_L}{\mu}\right) - \underline{x}\right)\right) = 1 - \left(\frac{\pi D_L}{\mu}\right)^{-\lambda\mu} \exp\left(\lambda \underline{x}\right).$$

The profit, $\underline{\pi}$, of the lowest quality-cost firm solves $F_{\Pi}(\underline{\pi}) = 0$, and thus verifies the expected property $\underline{\pi} = \frac{\mu}{D_L} \exp\left(\frac{\underline{x}}{\mu}\right)$. Inserting this value back into $F_{\Pi}(\pi)$ gives the expression in Proposition 6. The output distribution follows from the profit distribution:

$$F_Y(y) = \Pr\left(Y < y\right) = \Pr\left(\frac{\Pi}{\mu} < y\right) = F_{\Pi}(\mu y) = 1 - \left(\frac{\underline{\pi}}{\mu y}\right)^{\lambda\mu} = 1 - \left(\frac{\underline{y}}{y}\right)^{\lambda\mu},$$

where the lowest output, $\underline{y}$, is associated to the lowest profit, $\underline{\pi} = \mu \underline{y}$.

References

[1] Anderson, Chris (2006). The Long Tail: Why the Future of Business Is Selling Less of More. New York: Hyperion.

[2] Anderson, Simon and André de Palma (2000). From local to Global Competition, European Economic Review, 44, 423-448.

[3] Anderson, S. and A. de Palma (2001). Product Diversity in Asymmetric Oligopoly: Is the Quality of Consumer Goods too Low? Journal of Industrial Economics, 49, 113-135.

[4] Anderson, Simon P. and André de Palma (2015). Economic distributions and primitive distributions in monopolistic competition, CEPR Discussion Papers, 10748.

[5] Anderson, Simon P., André de Palma, and Jacques François Thisse (1992). Discrete Choice Theory of Product Differentiation. Cambridge, MA: MIT Press.

[6] Atkin, David, Azam Chaudhry, Shamyla Chaudhry, Amit K. Khandelwal, and Eric Verhoogen (2015). Markup and Cost Dispersion across Firms: Direct Evidence from Producer Surveys in Pakistan. American Economic Review, 105 (5): 537-44.

[7] Axtell, Rob (2001). Zipf distribution of U.S. firm sizes. Science 293, 1818–20.

[8] Baldwin, Richard, and James Harrigan (2011). Zeros, Quality, and Space: Trade Theory and Trade Evidence. American Economic Journal: Microeconomics 3(2), 60-88.

[9] Bertoletti, Paolo, and Federico Etro (2017). Monopolistic competition when income matters. Economic Journal, 127, 1217-1243.

[10] Blonigen, Bruce A. and Anson Soderbery (2010). Measuring the Benefits of Product Variety with an Accurate Variety Set. Journal of International Economics, 82, 168-180.

[11] Broda, Christian and David E. Weinstein (2006). Globalization and the Gains from Variety. Quarterly Journal of Economics, 121(2), 541-585.

[12] Cabral, Luis , and Jose Mata (2003). "On the evolution of the firm size distribution: Facts and theory." American Economic Review, 93(4), 1075-1090.

[13] Caplin, Andrew, and Barry Nalebuff (1991). "Aggregation and imperfect competition: On the existence of equilibrium." Econometrica, 59(1), 25–59.

[14] Chamberlin, Edward (1933). Theory of Monopolistic Competition. Cambridge, MA: Harvard University Press

[15] Coad, Alex (2009). On the distribution of product price and quality. Journal of Evolutionary Economics, 19(4), 589–604.

[16] Crozet, Matthieu, Keith Head, and Thierry Mayer (2011). Quality Sorting and Trade Firm-level Evidence for French Wine. Review of Economic Studies, 79(2), 609-644.

[17] di Giovanni, A. Levchenko, and Isabelle Méjean (2014). Firms, Destinations, and Aggregate Fluctuations. Econometrica, 82-4, 1303-1340

[18] Dixit, Avinash K., and Joseph S. Stiglitz (1977). Monopolistic Competition and Optimal Product Diversity. American Economic Review, 67(3), 297-308.

[19] Dhingra, Swati and John Morrow (2017). Monopolistic competition and optimum product diversity under firm heterogeneity, Journal of Political Economy, forthcoming.

[20] Fajgelbaum, Pablo D., Gene M. Grossman, and Elhanan Helpman (2011). Income Distribution, Product Quality, and International Trade. Journal of Political Economy, 119(4), 721-765.

[21] Feenstra, Robert C. (1994). New Product Varieties and the Measurement of International Prices, American Economic Review, 84(1), 157-177.

[22] Feenstra, Robert C. and John Romalis (2014). International Prices and Endogenous Quality, Quarterly Journal of Economics, 129(2), 477-527.

[23] Gabaix, Xavier (2016). Power Laws in Economics: An Introduction. Journal of Economic Perspectives. 30 (1), 185-206.

[24] Head, Keith, Thierry Mayer, and Mathias Thoenig (2014). Welfare and Trade without Pareto. American Economic Review 104(5): 310-16.

[25] Helpman, Elhanan and Paul R. Krugman (1985). Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy. MIT Press, Cambridge, Mass.

[26] Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple (2004). Export Versus FDI with Heterogeneous Firms. American Economic Review, 94(1): 300-316.

[27] Hitsch, Günter J., Ali Hortacsu, and Xiliang Lin (2017). Prices and Promotions in U.S. Retail Markets: Evidence from Big Data. Booth Discussion Paper 17-18, University of Chicago.

[28] Kaplan, Greg and Guido Menzio (2015). The Morphology Of Price Dispersion. International Economic Review, 56, 1165-1206.

[29] Kleiber, C. (2014). The generalized lognormal distribution and the Stieltjes moment problem. Journal of Theoretical Probability, 27(4), 1167-1177.

[30] Kugler, Maurice and Eric Verhoogen (2012). Prices, Plant Size, and Product Quality. Review of Economic Studies, 79(1): 307-339.

[31] Mrázová, Monika, J. Peter Neary, and Mathieu Parenti (2017). Sales and Markup Dispersion: Theory and Empirics. CEPR Discussion Paper 12044.

[32] McFadden, Daniel (1978). Modelling the Choice of Residential Location. in A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull (eds.), Spatial Interaction Theory and Planning Models, Amsterdam: North Holland, 75-96, Reprinted in J. Quigley (ed.), 1997, The economics of housing, Vol. I, Edward Elgar: London, 531-552.

[33] Melitz, Marc J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica, 71(6), 1695-1725.

[34] Nigai, Sergey (2017). A tale of two tails: Productivity distribution and the gains from trade. Journal of International Economics, 104, 44-62.

[35] Nocke, Volker and Stephen Yeaple (2008). Firm heterogeneity and the structure of US multinational activity. Review of Economic Studies, 75(2), 529-557.

[36] Pareto, Vilfredo (1965). La Courbe de la Répartition de la Richesse (Original in 1896): Oeuvres Complètes de Vilfredo. Ed. Busino G. Pareto. Geneva: Librairie Droz, 1-5.

[37] Zhelobodko, Evgeny, Sergey Kokovin, Mathieu Parenti, and Jacques-François Thisse (2012). Monopolistic competition: Beyond the constant elasticity of substitution. Econometrica 80(6), 2765-2784.