Online Appendix to:

"A Microfoundation for Normalized CES Production Functions with Factor–Augmenting Technical Change"

Jakub Growiec*

February 20, 2013

1 A microfoundation for the Cobb–Douglas aggregate production function

The "endogeneous technology choice" framework presented in Section 2 of the main text can also be used to derive the aggregate Cobb–Douglas production function (cf. Jones, 2005). The key change in assumptions that is required to produce such a result relates to the distribution of capital- and labor-augmenting ideas – which ought to be independently Pareto-distributed – and thus the shape of the technology menu; everything else is preserved. In Section 3 of the main text we have argued this to be empirically problematic, so that the aggregate CES production function (with gross complementarity of inputs) should in fact be considered a more plausible alternative. Nevertheless, the Cobb–Douglas case remains a useful benchmark for comparisons because it is so frequently used in the literature. Let us address it in the current section of the online appendix.

*Warsaw School of Economics, Institute of Econometrics, al. Niepodległości 162, 02-554 Warszawa, Poland, and National Bank of Poland, Economic Institute. E-mail: jakub.growiec@sgh.waw.pl.

1.1 Modification of the framework

Let us now replace Assumption 2 from the main text with the following one:

Assumption 1 (modification of Assumption 2) *The technology menu, defined in the (a,b) space, is given by the equality:*

$$H(a,b)=\left(\frac{a}{\lambda_a}\right)^{\phi_L}\left(\frac{b}{\lambda_b}\right)^{\phi_K}=N, \qquad \phi_K,\phi_L>0. \tag{1}$$

The shape of the technology menu given by equation (1) is consistent with the assumption that $\tilde{a}$ and $\tilde{b}$ are independently Pareto-distributed, with shape parameters ϕ_L and ϕ_K, respectively:

$$P(\tilde{a}>a)=\left(\frac{\lambda_a}{a}\right)^{\phi_L}, \qquad P(\tilde{b}>b)=\left(\frac{\lambda_b}{b}\right)^{\phi_K}, \tag{2}$$

for $a>\lambda_a$ and $b>\lambda_b$. In such a case, $N=\frac{1}{P(\tilde{a}>a,\tilde{b}>b)}>1$. Just like in the main text, we assume N to be fixed, and allow λ_a and λ_b to rise over time thanks to directed R&D.

The same functional form of the technology menu was assumed by Jones (2005), but with the unnecessary restriction of proportional (Hicks–neutral) augmentation of the technology menu, which is now relaxed.

1.2 The aggregation result

As in the CES case, deriving optimal technology choices from the firms' optimization problem is straightforward. Inserting these optimal choices into the LPF, we obtain the following aggregation result.

Proposition 1 *If Assumption 1 above as well as Assumptions 1 and 3 from the main text hold, then the aggregate production function takes the normalized Cobb–Douglas form:*

$$Y=Y_0\left(\frac{\lambda_a}{\lambda_{a0}}\right)^{\frac{\phi_L}{\phi_L+\phi_K}}\left(\frac{\lambda_b}{\lambda_{b0}}\right)^{\frac{\phi_K}{\phi_L+\phi_K}}\left(\frac{K}{K_0}\right)^{\frac{\phi_K}{\phi_L+\phi_K}}\left(\frac{L}{L_0}\right)^{\frac{\phi_L}{\phi_L+\phi_K}}. \tag{3}$$

Proof (and generalization to n inputs): see Section 3 of this online appendix. ■

The interpretation of the parameters of the aggregate Cobb–Douglas production function is the following:

- the distribution parameter of the aggregate Cobb–Douglas production function, equal to the capital's partial elasticity and the (constant) capital income share, takes the value $\pi_0 = \frac{r_0 K_0}{Y_0} = \frac{\phi_K}{\phi_L + \phi_K}$,

- partial elasticities of capital and labor in the aggregate production function are proportional to the shape parameters of the Pareto distributions of their respective factor-augmenting technologies and sum up to one (guaranteeing constant returns to scale),

- the multiplicative constant term is Y_0. Thanks to normalization, it is thus exactly equal to the multiplicative constant term of the LPF,

- the constant parameter N does not appear in the aggregate production function,[1]

- the capital-and labor-augmenting parameters of the technology menu, λ_b and λ_a respectively, enter the aggregate production function multiplicatively, taken to their respective powers ϕ_K and ϕ_L. Growth in aggregate output is thus invariant to the direction of R&D.

1.3 Direction of technical change vs. direction of R&D

It is also easily verified that under endogeneous technology choice, the Cobb–Douglas case provides very specific implications for the direction of technical change. Log-differentiating firms' optimal technology choices with respect to time and comparing terms we obtain:

$$\hat{a} = \hat{y} = \frac{\phi_L}{\phi_L + \phi_K}\hat{\lambda}_a + \frac{\phi_K}{\phi_L + \phi_K}\hat{\lambda}_b + \frac{\phi_K}{\phi_L + \phi_K}\hat{k}, \tag{4}$$

$$\hat{b} = \hat{y} - \hat{k} = \frac{\phi_L}{\phi_L + \phi_K}\hat{\lambda}_a + \frac{\phi_K}{\phi_L + \phi_K}\hat{\lambda}_b - \frac{\phi_L}{\phi_L + \phi_K}\hat{k}. \tag{5}$$

Hence, it follows that in the Cobb–Douglas case, no matter what the direction of R&D is, i.e., irrespective of the values of $\hat{\lambda}_a$ and $\hat{\lambda}_b$, firms will always adjust the labor-augmenting technology on one-to-one basis to changes in output per worker y,

[1] Again, one could easily reparametrize the technology menu, fixing either λ_a or λ_b and allowing N to vary across time. In such case, the ratio N/N_0 (which now drops out) would appear in equation (3).

and capital-augmenting technology will be, accordingly, always adjusted one-to-one to changes in output per unit of capital y/k. Hence, as shown by Jones (2005), technological change must be purely labor-augmenting along the balanced growth path, where the output–capital ratio y/k is constant.

Assuming that factors of production are remunerated according to their marginal products, the capital income share is now fixed at $\pi_0 = \frac{\phi_K}{\phi_L+\phi_K}$, and the labor income share is fixed at $1-\pi_0 = \frac{\phi_L}{\phi_L+\phi_K}$, for all times t.

2 A microfoundation for Pareto UFP distributions

As discussed in the main text, the robustness of the result of Weibull-distributed UFPs stems from the fact that the Weibull distribution is, under very general conditions, the limiting distribution for sample minima. The Pareto distribution cannot be derived that way. The latter distribution is, on the other hand, closely related to the limiting distribution of exceedances of random variables above a given threshold – the so-called Generalized Pareto distribution. Thus, after some reparametrization, the Pareto distribution can also be derived as a robust limiting distribution; this requires a substantial modification of the assumptions on the underlying R&D process, though.

To develop this argument, we shall return to the original Jones's (2005) view that the technology menu is a convex hull of a number of *simple* ideas (consisting of a single component only), distributed according to some underlying distribution $\mathcal{F}$. (We shall tentatively ignore all the arguments against such standpoint which we put forward in Section 3 of the main text.) Furthermore, we shall reconcile this view with Assumption 5, assuming that z_a and z_b are sufficiently small. As it turns out, taking these two premises together leads – by the force of extreme value theory – to a surprisingly robust conclusion: it is found that in such case the individual UFPs – modeled as distributions of *exceedances over a sufficiently high threshold* – must follow the Generalized Pareto distribution. This distribution has a cdf equal to $1-(1+\xi x)^{-1/\xi}$ and simplifies to the Pareto distribution after appropriate reparametrization. Hence, the Jones' view of the technology choice process, dwelling on substitutability between researchers and an increasing quality threshold for new ideas, rather than complementarity across components of ideas and their ever increasing complexity, is consistent with the Cobb–Douglas aggregate production function. However, this view contradicts the empirical findings of increasing technological complexity of new inventions, skill-biased technical change, mounting burden of knowledge, and increasing R&D collaboration.

More formally, we make the following assumption:

Assumption 2 (modification of Assumption 4) *The (capital- or labor-augmenting) R&D sector consists of an infinity of researchers located along the unit interval $I = [0, 1]$. At each instant t, every researcher $i \in I$ determines the quality of her innovation ($\tilde{b}_i$ or $\tilde{a}_i$, respectively) by taking a single independent draw from the elementary idea distribution $\mathcal{F}$. The distribution $\mathcal{F}$ has positive density on $[w, +\infty)$, and zero density otherwise, and satisfies the condition of a regularly varying upper tail:*

$$\lim_{p \to +\infty} \frac{1 - \mathcal{F}(w + px)}{1 - \mathcal{F}(w + p)} = x^{-\frac{1}{\xi}} \tag{6}$$

for all $x > 0$ and a certain $\xi > 0$.

We are interested in the distribution of UFPs $\tilde{b}_i$ and $\tilde{a}_i$ conditional on exceeding a sufficiently high threshold, $\bar{b}$ or $\bar{a}$, respectively. To economize on space, let us concentrate on the case of capital-augmenting ideas b for the rest of this section, as the results regarding labor-augmenting ideas a are entirely symmetric.

We shall make Assumption 5 once again, and consistently denote the fraction of researchers whose UFP draws have exceeded the given threshold $\bar{b}$ as $z_{\bar{b}} = P(\tilde{b} > \bar{b})$. The distribution of exceedances of $\tilde{b}$ with cdf $\mathcal{F}$ over the threshold $\bar{b} > w$ (captured by the probability $P(\tilde{b} > b | \tilde{b} > \bar{b})$ as a function of b) is then determined by the form of the particular cdf $\mathcal{F}$ as well as the value of $\bar{b}$ (or equivalently $z_{\bar{b}}$). In the limit where $\bar{b} \to \infty$, or equivalently $z_{\bar{b}} \to 0$, however, the following limit result obtains:

Proposition 2 *If Assumption 2 holds, then there exists a positive measurable function $\beta : \mathbb{R} \to \mathbb{R}$ satisfying*

$$\lim_{x \to \infty} \frac{\beta(x)}{x} = \xi, \tag{7}$$

such that as $z_{\bar{b}} \to 0$ (and thus $\bar{b} \to \infty$), the distribution of any sequence of random variables X defined via

$$P(X > x) = P(\tilde{b} > b | \tilde{b} > \bar{b}) \tag{8}$$

converges in distribution to the Generalized Pareto distribution with the shape parameter $\xi > 0$*:*

$$\frac{1-\mathcal{F}\left(\bar{b}+x\beta(\bar{b})\right)}{1-\mathcal{F}(\bar{b})} \xrightarrow{d} (1+\xi x)^{-1/\xi}. \tag{9}$$

Proof. The proposition follows directly from the Pickands–Balkema–de Haan extreme value theorem (Balkema and de Haan, 1974; Pickands, 1975), applied to the distribution $\mathcal{F}$ which has a regularly varying upper tail with $\xi > 0$. ■

Hence, for sufficiently small $z_{\bar{b}}$, the distribution of exceedances of $\tilde{b}$ over the given threshold $\bar{b}$ can be arbitrarily well approximated by the Generalized Pareto distribution regardless of the underlying distribution $\mathcal{F}$. Moreover, a specific change of variables given by $\breve{b} = \lambda_b(1+\xi X)$ can be used to redirect our analysis to the case of pure Pareto distributions:

Proposition 3 *If the (non-negative) random variable* X *has a Generalized Pareto distribution with the parameter* $\xi > 0$*, then we can define* $\breve{b} = \lambda_b(1+\xi X)$*, with values* $\breve{b} > \lambda_b$*, such that* $\breve{b}$ *has a Pareto distribution with the slope parameter* $1/\xi$*.*

Proof.

$$P(\breve{b} > b) = P(\lambda_b(1+\xi X) > b) = P\left(X > \frac{b-\lambda_b}{\xi\lambda_b}\right) = \left(1+\xi\frac{b-\lambda_b}{\xi\lambda_b}\right)^{-1/\xi} = \left(\frac{\lambda_b}{b}\right)^{1/\xi}. \blacksquare$$

Thus, having computed the value of ξ from the underlying distribution $\mathcal{F}$ according to equation (6), and assuming independence of $\breve{b}$ and $\breve{a}$, we can construct the technology menu as a product of contour lines of two Pareto distributions, getting us directly to Assumption 1 above. As a side remark, please note the identity $1/\xi = \phi_K$, signifying that the capital share of the aggregate production function is fundamentally determined by the tail properties of the underlying idea distribution $\mathcal{F}$.

3 Generalization of the aggregation procedure to n inputs

As announced in the main text, all our results go through in the general case of n-input production functions as well. Let us now discuss this case.

3.1 The normalized CES case

First, let us show that if ideas (UFPs), augmenting each of the n production inputs, are *independently Weibull-distributed* (and the LPFs are normalized CES or Leontief functions), then the resultant aggregate production is normalized CES as well. To this end, we shall use the following generalized assumptions. By x_i, $i = 1, 2, ..., n$ we shall denote the inputs, and by a_i, $i = 1, 2, ..., n$ – unit factor productivities.

Assumption 3 *The n-input local production function (LPF) takes either the normalized CES or the normalized Leontief form:*

$$Y = \begin{cases} Y_0 \left(\sum_{i=1}^{n} \pi_{0i} \left(\frac{a_i x_i}{a_{0i} x_{0i}} \right)^{\theta} \right)^{\frac{1}{\theta}}, & \text{if } \sigma_{LPF} \in (0,1), \\ Y_0 \min_{i=1,\dots,n} \left\{ \left(\frac{a_i x_i}{a_{0i} x_{0i}} \right) \right\}, & \text{if } \sigma_{LPF} = 0, \end{cases} \tag{10}$$

where $\theta \in [-\infty, 0)$ is the substitutability parameter*, related to the elasticity of substitution along the LPF via $\sigma_{LPF} = \frac{1}{1-\theta}$. The Leontief LPF, with $\sigma_{LPF} = 0$, is obtained as a special case of the more general normalized CES class of LPFs by taking the limit $\theta \to -\infty$ (we denote this case as $\theta = -\infty$ for simplicity). π_{0i} is the income share of i-th factor at t_0. Factor income shares sum up to unity:*

$$\sum_{i=1}^{n} \pi_{0i} = 1, \tag{11}$$

and the LPF exhibits constant returns to scale.

Assumption 4 *The technology menu, specified in the $(a_1, ..., a_n)$ space, is given by the equality:*

$$H(a_1, ..., a_n) = \sum_{i=1}^{n} \left(\frac{a_i}{\lambda_{ai}} \right)^{\alpha} = N, \qquad \lambda_{a1}, ..., \lambda_{an}, \alpha, N > 0. \tag{12}$$

The technology menu is understood as a contour line of the cumulative distribution function of the joint n-variate distribution of factor-augmenting ideas $\tilde{a}_i$, $i = 1, ..., n$. Under independence of the n dimensions (so that marginal distributions are multiplied by one another), equation (12) obtains if and only if the marginal distributions are Weibull with the same shape parameter $\alpha > 0$ (Growiec, 2008b):

$$P(\tilde{a}_i > a_i) = e^{-\left(\frac{a_i}{\lambda_{ai}}\right)^{\alpha}}, \qquad i = 1, 2, ..., n, \tag{13}$$

where all $a_i > 0$. Under such a parametrization, we have

$$P(\tilde{a}_1 > a_1, ..., \tilde{a}_n > a_n) = e^{-\sum_{i=1}^{n}\left(\frac{a_i}{\lambda_{ai}}\right)^{\alpha}}, \tag{14}$$

and thus the parameter N in equation (12) is interpreted as $N = -\ln P(\tilde{a}_1 > a_1, ..., \tilde{a}_n > a_n) > 0$.

The case where $\tilde{a}_i, i = 1, ..., n$ are independently Pareto-distributed leads to a different specification of the technology menu and will be considered separately in the next subsection. If they are Weibull-distributed but dependent, or independent but following some other distribution than Pareto or Weibull, the resultant aggregate production does not belong to the CES class and will not be considered here.

Assumption 5 *Firms choose the technology n-tuple $(a_1, ..., a_n)$ optimally, subject to the current technology menu, such that their profit is maximized:*

$$\max_{a,b} \left\{ Y_0 \left(\sum_{i=1}^{n} \pi_{0i} \left(\frac{a_i x_i}{a_{0i} x_{0i}} \right)^{\theta} \right)^{\frac{1}{\theta}} \right\} \qquad s.t. \qquad \sum_{i=1}^{n} \left(\frac{a_i}{\lambda_{ai}} \right)^{\alpha} = N. \tag{15}$$

Factor remuneration, taken into account in the firms' profit maximization problem, does not depend on the chosen technology so it can be safely omitted from the above optimization problem.[2]

Finally, second order conditions require us to assume that $\alpha > \theta$, so that the interior stationary point of the above optimization problem is a maximum. For the resultant aggregate production function to be concave with respect to x_i, $i = 1, ..., n$, we need to assume furthermore that $\alpha - \theta - \alpha\theta > 0$. All these conditions are satisfied automatically in the case $\alpha > 0 > \theta$, on which we concentrate here. The inputs are gross complements along the aggregate production function.

The framework provides direct results on the firm's optimal technology choice. First, at time t_0, when $Y = Y_0$ and $x_i = x_{0i}, \lambda_{ai} = \lambda_{a0i}$ is assumed for all $i = 1, ..., n$, the optimal technology choice satisfies:

$$a_{0i}^{*} = (N\pi_{0i})^{\frac{1}{\alpha}} \lambda_{a0i}, \qquad i = 1, ..., n, \tag{16}$$

[2] In the case of Leontief LPFs, optimization implies $\frac{a_i x_i}{a_{0i} x_{0i}} = \frac{a_j x_j}{a_{0j} x_{0j}}$ for all $i, j = 1, ..., n$.

where λ_{a0i} is the value of λ_{ai} at time t_0. Values of a_{0i}^* will be used as a_{0i} in the normalization at the local level in all subsequent derivations.

For any other moment in time $t \neq t_0$, the optimal technology choices are:

$$\left(\frac{a_j}{a_{0j}}\right)^* = \frac{\lambda_{aj}}{\lambda_{a0j}} \left(\sum_{i=1}^{n} \pi_{0i} \left(\frac{\lambda_{ai}}{\lambda_{aj}} \frac{\lambda_{a0j}}{\lambda_{a0i}} \frac{x_i x_{0j}}{x_{0i} x_j}\right)^{\frac{\alpha\theta}{\alpha-\theta}}\right)^{-\frac{1}{\alpha}}, \tag{17}$$

for all $j = 1, ..., n$, where $\frac{\alpha\theta}{\alpha-\theta}$ is substituted with $-\alpha$ in the case of Leontief LPFs ($\theta = -\infty$).

Inserting these optimal technology choices into the LPF, we obtain the following aggregation result.

Proposition 4 *If Assumptions 3-5 hold, then the aggregate production function takes the normalized CES form:*

$$Y = Y_0 \left(\sum_{i=1}^{n} \pi_{0i} \left(\frac{\lambda_{ai}}{\lambda_{a0i}} \frac{x_i}{x_{0i}}\right)^{\frac{\alpha\theta}{\alpha-\theta}}\right)^{\frac{\alpha-\theta}{\alpha\theta}}. \tag{18}$$

Again, $\frac{\alpha\theta}{\alpha-\theta}$ is substituted with $-\alpha$ in the case of Leontief LPFs. Hence, the normalized CES result obtains both in the case of CES and Leontief LPFs.

Proof. First, in the case of CES LPFs, we form the Lagrangean:

$$\mathcal{L} = Y_0 \left(\sum_{i=1}^{n} \pi_{0i} \left(\frac{a_i x_i}{a_{0i} x_{0i}}\right)^{\theta}\right)^{\frac{1}{\theta}} + \Lambda \cdot \left\{\sum_{i=1}^{n} \left(\frac{a_i}{\lambda_{ai}}\right)^{\alpha} - N\right\}. \tag{19}$$

Differentiating it with respect to a_i, $i = 1, ..., n$, and substituting for Λ yields:

$$\left(\frac{a_i}{a_j}\right)^{\alpha-\theta} = \frac{\pi_{0i}}{\pi_{0j}} \left(\frac{\lambda_{ai}}{\lambda_{aj}}\right)^{\alpha} \left(\frac{x_i}{x_j} \frac{a_{0j} x_{0j}}{a_{0i} x_{0i}}\right)^{\theta}, \tag{20}$$

for all $i, j = 1, ..., n$. Considering first the reference point of time t_0, when $x_i = x_{0i}, \lambda_{ai} = \lambda_{a0i}, a_i = a_{0i}$ for all $i = 1, ..., n$, we obtain:

$$\frac{a_{0i}}{a_{0j}} = \left(\frac{\pi_{0i}}{\pi_{0j}}\right)^{\frac{1}{\alpha}} \frac{\lambda_{a0i}}{\lambda_{a0j}}. \tag{21}$$

Using the specification of the technology menu (12) as well as the assumption that $\sum_{i=1}^{n} \pi_{0i} = 1$, we obtain:

$$a_{0i}^* = (N\pi_{0i})^{\frac{1}{\alpha}} \lambda_{a0i}, \qquad i = 1, ..., n. \tag{22}$$

For $t \neq t_0$, by plugging (22) into (20), using (12) again and rearranging, we obtain that:

$$\left(\frac{a_j}{a_{0j}}\right)^* = \frac{\lambda_{aj}}{\lambda_{a0j}} \left(\sum_{i=1}^{n} \pi_{0i} \left(\frac{\lambda_{ai}}{\lambda_{aj}} \frac{\lambda_{a0j}}{\lambda_{a0i}} \frac{x_i x_{0j}}{x_{0i} x_j}\right)^{\frac{\alpha\theta}{\alpha-\theta}}\right)^{-\frac{1}{\alpha}}, \tag{23}$$

for all $j = 1, ..., n$.

Plugging this into the LPF (10) and rearranging, we obtain the final result.

Given our parametric assumptions, second-order conditions for the maximization of the Lagrangean hold. To demonstrate this, it is useful to note that maximizing $\mathcal{L}$ is equivalent to minimizing the following transformed Lagrangean $\mathcal{L}_{min}$ (where the maximand function is taken to the power $\theta < 0$ for simplicity):

$$\mathcal{L}_{min} = Y_0^{\theta} \sum_{i=1}^{n} \pi_{0i} \left(\frac{a_i x_i}{a_{0i} x_{0i}}\right)^{\theta} + \Lambda_{min} \cdot \left\{\sum_{i=1}^{n} \left(\frac{a_i}{\lambda_{ai}}\right)^{\alpha} - N\right\}. \tag{24}$$

We obtain the following second-order derivatives of $\mathcal{L}_{min}$ (after inserting the first order condition to get rid of Λ_{min}):

$$\frac{\partial^2 \mathcal{L}_{min}}{\partial a_i^2} = \theta(\theta - \alpha) Y_0^{\theta} \pi_{0i} \left(\frac{a_i x_i}{a_{0i} x_{0i}}\right)^{\theta} \frac{1}{a_i^2} > 0, \tag{25}$$

$$\frac{\partial^2 \mathcal{L}_{min}}{\partial a_i \partial a_j} = 0, \tag{26}$$

and thus $\mathcal{L}_{min}$ is minimized.

In the case of Leontief LPFs, instead of forming the Lagrangean, one should use the equality $\frac{a_i x_i}{a_{0i} x_{0i}} = \frac{a_j x_j}{a_{0j} x_{0j}}$ for all $i, j = 1, ..., n$ – which must hold because of the assumption that the representative firm maximizes profits. Since equations (12) and (22) still hold, plugging these equalities into the LPF yields

$$Y = Y_0 \frac{a_1 x_1}{a_{01} x_{01}} = Y_0 \left(\sum_{i=1}^{n} \pi_{0i} \left(\frac{\lambda_{ai}}{\lambda_{a0i}} \frac{x_i}{x_{0i}}\right)^{-\alpha}\right)^{-\frac{1}{\alpha}}. \tag{27}$$

Please note that the same result is obtained by taking the case of CES LPFs and considering the limit $\theta \to -\infty$. ■

The corollary on factor income shares goes through in the n-dimensional case as well:

Corollary 1 *Assuming that factors are priced at their marginal product, the factor income shares are equal to:*

$$\pi_i = \frac{\pi_{0i}\left(\frac{\lambda_{ai}}{\lambda_{a0i}}\frac{x_i}{x_{0i}}\right)^{\frac{\alpha\theta}{\alpha-\theta}}}{\sum_{i=1}^n \pi_{0i}\left(\frac{\lambda_{ai}}{\lambda_{a0i}}\frac{x_i}{x_{0i}}\right)^{\frac{\alpha\theta}{\alpha-\theta}}}, \qquad i = 1, ..., n. \tag{28}$$

3.2 The Cobb–Douglas case

Let us now replace Assumption 4 with the following one:

Assumption 6 (modification of Assumption 4) *The technology menu, specified in the $(a_1, ..., a_n)$ space, is given by the equality:*

$$H(a_1, ..., a_n) = \prod_{i=1}^n \left(\frac{a_i}{\lambda_{ai}}\right)^{\phi_i} = N, \qquad \phi_i > 0, i = 1, ..., n. \tag{29}$$

The current shape of the technology menu is consistent with the assumption that $\tilde{a}_i$'s are independently Pareto-distributed, with respective shape parameters ϕ_i. In such case, $N = \frac{1}{P(\tilde{a}_1 > a_1, ..., \tilde{a}_n > a_n)} > 1$.

At t_0, when $Y = Y_0$ and $x_i = x_{0i}, \lambda_{ai} = \lambda_{a0i}$ is assumed for $i = 1, ..., n$, the optimal choice is indeterminate, provided that

$$\pi_{0i} = \frac{\phi_i}{\sum_{i=1}^n \phi_i}, \qquad i = 1, ..., n. \tag{30}$$

This restriction means that the factor income shares should be equal to $\frac{\phi_i}{\sum_{i=1}^n \phi_i}$. Thus, $\pi_{02}, ..., \pi_{0n}$ cease to be free parameters, and $a_{02}, ..., a_{0n}$ become free parameters instead (the remaining technology choice a_{01} is then calculated according to the technology menu).

At any other moment in time $t \neq t_0$, and given $a_{0i}, i = 1, ..., n$, the optimal technology choices are:

$$\left(\frac{a_i}{a_{0i}}\right)^* = \frac{x_{0i}}{x_i}\prod_{i=1}^n \left(\frac{\lambda_{ai}}{\lambda_{a0i}}\frac{x_i}{x_{0i}}\right)^{\frac{\phi_i}{\sum_{i=1}^n \phi_i}}, \qquad i = 1, ..., n. \tag{31}$$

Inserting these optimal technology choices into the LPF, we obtain the following result.

Proposition 5 *If Assumptions 7, 9, and 10 hold, then the aggregate production function takes the normalized Cobb–Douglas form:*

$$Y = Y_0 \prod_{i=1}^{n} \left(\frac{\lambda_{ai}}{\lambda_{a0i}} \frac{x_i}{x_{0i}} \right)^{\frac{\phi_i}{\sum_{i=1}^{n} \phi_i}}. \tag{32}$$

Proof. We form the Lagrangean:

$$\mathcal{L} = Y_0 \left(\sum_{i=1}^{n} \pi_{0i} \left(\frac{a_i x_i}{a_{0i} x_{0i}} \right)^{\theta} \right)^{\frac{1}{\theta}} + \Lambda \cdot \left\{ \prod_{i=1}^{n} \left(\frac{a_i}{\lambda_{ai}} \right)^{\phi_i} - N \right\}. \tag{33}$$

Differentiating it with respect to a_i, $i = 1, ..., n$, and substituting for Λ yields:

$$\left(\frac{a_i x_i}{a_j x_j} \frac{a_{0j} x_{0j}}{a_{0i} x_{0i}} \right)^{\theta} \frac{\pi_{0i}}{\pi_{0j}} \frac{\phi_j}{\phi_i} = 1, \tag{34}$$

for all $i, j = 1, ..., n$. Considering first the reference point of time t_0, when $x_i = x_{0i}, \lambda_{ai} = \lambda_{a0i}, a_i = a_{0i}$ for all $i = 1, ..., n$, we obtain:

$$\frac{\pi_{0i}}{\pi_{0j}} = \frac{\phi_i}{\phi_j}. \tag{35}$$

Using the assumption that $\sum_{i=1}^{n} \pi_{0i} = 1$, we obtain that at t_0, optimal technology choice is indeterminate provided that:

$$\pi_{0i} = \frac{\phi_i}{\sum_{i=1}^{n} \phi_i}, \qquad i = 1, ..., n. \tag{36}$$

For $t \neq t_0$, by plugging (35) into (34), using (29) and rearranging, we obtain that:

$$\left(\frac{a_i}{a_{0i}} \right)^{*} = \frac{x_{0i}}{x_i} \prod_{i=1}^{n} \left(\frac{\lambda_{ai}}{\lambda_{a0i}} \frac{x_i}{x_{0i}} \right)^{\frac{\phi_i}{\sum_{i=1}^{n} \phi_i}}, \qquad i = 1, ..., n. \tag{37}$$

for all $j = 1, ..., n$.

Plugging this into the LPF (10) and rearranging, we obtain the final result.

Given our parametric assumptions, second-order conditions for the maximization of the Lagrangean hold. To prove this, it is useful to note that maximizing $\mathcal{L}$ is equivalent to minimizing the following transformed Lagrangean $\mathcal{L}_{min}$ (where, for simplicity, the maximand function is taken to the power $\theta < 0$ and a log-transformation is applied to the restriction):

$$\mathcal{L}_{min} = Y_0^{\theta} \sum_{i=1}^{n} \pi_{0i} \left(\frac{a_i x_i}{a_{0i} x_{0i}} \right)^{\theta} + \Lambda_{min} \cdot \left\{ \sum_{i=1}^{n} \phi_i (\ln a_i - \ln \lambda_{ai}) - \ln N \right\}. \tag{38}$$

We obtain the following second-order derivatives of $\mathcal{L}_{min}$ (after inserting the first order condition to get rid of Λ_{min}):

$$\frac{\partial^2 \mathcal{L}_{min}}{\partial a_i^2} = \theta^2 Y_0^\theta \pi_{0i} \left(\frac{a_i x_i}{a_{0i} x_{0i}} \right)^\theta \frac{1}{a_i^2} > 0, \tag{39}$$

$$\frac{\partial^2 \mathcal{L}_{min}}{\partial a_i \partial a_j} = 0, \tag{40}$$

and thus $\mathcal{L}_{min}$ is minimized.

In the case of Leontief LPFs, instead of forming the Lagrangean, one should use the equality $\frac{a_i x_i}{a_{0i} x_{0i}} = \frac{a_j x_j}{a_{0j} x_{0j}}$ for all $i, j = 1, ..., n$ – which must hold because of the assumption that the representative firm maximizes profits. Since equation (29) still holds, plugging these equalities into the LPF yields

$$Y = Y_0 \frac{a_1 x_1}{a_{01} x_{01}} = Y_0 \prod_{i=1}^{n} \left(\frac{\lambda_{ai}}{\lambda_{a0i}} \frac{x_i}{x_{0i}} \right)^{\frac{\phi_i}{\sum_{i=1}^{n} \phi_i}} . \blacksquare \tag{41}$$

4 A reinterpretation of the Cobb–Douglas case in terms of technology adoption costs

As apparent from the related contribution of León-Ledesma and Satchi (2011), the variant of the current "endogeneous technology choice" framework which leads to the Cobb–Douglas result (presented in this appendix), could also be reinterpreted in terms of (Hicks-neutral) technology adoption costs. Namely, as posited by these authors, one could assume that the LPF takes the following specific *unnormalized* CES form (see also Growiec, 2008a):

$$Y = \Gamma f(\eta) \left(\eta \left(\lambda_b K \right)^\theta + (1 - \eta) \left(\lambda_a L \right)^\theta \right)^{\frac{1}{\theta}}, \qquad \theta < 0, \Gamma > 0, \eta \in [0, 1], \tag{42}$$

where the assumption $\theta < 0$ mirrors gross complementarity of inputs along the LPF. The inclusion of the Hicks-neutral term $f(\eta)$ in the local production function is meant to capture adoption costs in the production process. As reflected by the posited inverse U-shape of $f(\eta)$, this specification implies that highly labor- or capital-intensive technologies are costly, whereas "intermediate" technologies, using both factors in moderation, are relatively cheap. All technology adoption costs are borne in the form of Hicks-neutral technical inefficiency.

Furthermore, assuming *unit invariance* of the LPF – that is, requiring that the functional form $f(\eta)$ does not change with the units of labor measurement (e.g., hours worked, number of full-time equivalent employees, etc.) – León-Ledesma and Satchi (2011) obtain that

$$f(\eta) = (\eta^{\gamma}(1-\eta)^{1-\gamma})^{-\frac{1}{\theta}}. \tag{43}$$

Upon maximization of eq. (42) with respect to $\eta \in [0,1]$, it is obtained that the optimal choice of η satisfies:

$$\frac{1-\eta}{\eta} = \frac{1-\gamma}{\gamma}\left(\frac{\lambda_b K}{\lambda_a L}\right)^{\theta}. \tag{44}$$

Inserting this optimal choice into (42), the aggregate production function is derived as:

$$Y = \Gamma\left(\gamma^{\gamma}(1-\gamma)^{1-\gamma}\right)^{-\frac{1}{\theta}}(\lambda_b K)^{\gamma}(\lambda_a L)^{1-\gamma}. \tag{45}$$

Hence, the aggregate production function is Cobb–Douglas, just like in Proposition 1. In consequence, it follows that the assumption of Hicks-neutral technology adoption costs, coupled with unit invariance, is equivalent to assuming that the technology menu takes the form (1), consistent with the assumption that UFPs are independently Pareto-distributed.

Two caveats remain when discussing this analogy, though. First, the adoption cost mechanism proposed by León-Ledesma and Satchi (2011), despite its analytical simplicity and intuitive appeal, does not take normalization of CES functions into account. Under normalization, however, the CES production function is itself invariant to the choice of units of measurement of capital and labor, and thus imposing unit invariance on top of that does not place any further restrictions on the functional form of $f(\eta)$, rendering the analytical assumption (43) unmotivated. This implication could potentially generate a wider variety of aggregate production functions that could be derived using the current "adoption costs" framework once the restriction (43) is relaxed. Second, the assumption that technology adoption costs are borne in the form of Hicks-neutral technical ineffiency is very likely to play a role in generating the multiplicative (Cobb–Douglas) form of the aggregate production function, too. If these costs were borne, for instance, in the form of factor-specific UFP losses, then the Cobb–Douglas result would likely fail, just as it fails in our current setup if the technology menu takes a different form than the one required by Assumption 1 above.

References

[1] Balkema, A., L. de Haan (1974), "Residual Life Time at Great Age", *Annals of Probability* 2, 792–804.

[2] Growiec, J. (2008a), "A New Class of Production Functions and an Argument Against Purely Labor-Augmenting Technical Change", *International Journal of Economic Theory* 4(4), 483-502.

[3] Growiec, J. (2008b), "Production Functions and Distributions of Unit Factor Productivities: Uncovering the Link", *Economics Letters* 101(1), 87-90.

[4] Jones, C.I. (2005), "The Shape of Aggregate Production Functions and the Direction of Technical Change", *Quarterly Journal of Economics* 120(2), 517-549.

[5] León-Ledesma, M.A., M. Satchi (2011), "The Choice of CES Production Techniques and Balanced Growth", University of Kent, School of Economics DP 11/13.

[6] Pickands, J. (1975), "Statistical Inference Using Extreme Order Statistics", *Annals of Statistics* 3, 119–131.