Bucci, Alberto; Ushchev, Philip

Conference Paper

The sources of sharing externalities: Specialization vs Competition

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Bucci, Alberto; Ushchev, Philip (2015) : The sources of sharing externalities: Specialization vs Competition, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/124811

The Sources of Sharing Externalities: Specialization versus Competition*

Alberto Bucci† Philip Ushchev‡

December 12, 2014

Abstract

We explore the nature of *"sharing externalities"* in a non-CES model of monopolistic competition with a differentiated intermediate good. We find that, in addition to the effect that an increase in the number of available varieties of intermediate inputs may have on the total factor productivity of the final good sector (specialization/complexity effect), another important effect comes to play a prominent role in shaping market outcomes. This second effect is due to market interactions between producers of intermediate inputs (competition effect). These two effects may work either in the same or in opposing directions, depending on how the elasticity of technological substitution across intermediate inputs varies with the number of inputs (number of firms). The competition effect vanishes in the limiting case in which the technology employed in the final good sector is a traditional CES.

Keywords: Sharing Externalities; External Increasing Returns to Scale; Variable Elasticity of Substitution; Specialization Effect; Competition Effect

JEL Classification: D24, D43, F12, L13

*We owe special thanks to Frédéric Robert-Nicoud for careful reading an earlier version of the manuscript and providing a number of insightful comments. We also thank Kristian Behrens, Yasusada Murata, Dennis Novy, Mathieu Parenti, and Jacques-François Thisse for their useful comments and suggestions. This article was commenced during a visit of Alberto Bucci at the National Research University–Higher School of Economics, Center for Market Studies and Spatial Economics of Saint Petersburg. We are grateful to the Government of the Russian Federation for financial support (under the grant 11.G34.31.0059).

†University of Milan. Email: alberto.bucci@unimi.it

‡**Corresponding Author.** National Research University–Higher School of Economics, Center for Market Studies and Spatial Economics - Rimsky Korsakov avenue, 47, 199034 Saint Petersburg (Russian Federation). Email: ph.ushchev@gmail.com

1 Introduction

Since Adam Smith (1776), it is widely acknowledged that an efficient division of labor is a key driver of increasing returns and gains from specialization. However, explicit ways of modeling these gains have been developed solely in the last few decades. In fact, only in the wake of the path-breaking papers by Dixit and Stiglitz (1977), Ethier (1982), Krugman (1979; 1981) and Lancaster (1980), a huge number of models has been put forward where free entry is combined with product differentiation and various forms of specialization.

Ethier (1982), in particular, developed a model of monopolistic competition with intermediate goods, which later on has become a workhorse in endogenous growth theory with horizontal innovations and in agglomeration economics. The key feature of his model is the presence of *sharing externalities* (Duranton and Puga, 2004; Fujita and Thisse, 2013), which generate endogenous increasing returns to scale. More precisely, it is postulated that a larger market leads to a wider differentiation of intermediate inputs, which, in turn, results in deeper specialization and higher total factor productivity (TFP) in the final good sector. This is the *specialization effect*, which makes deeper product differentiation in the intermediate sector beneficial to the producers of the final good.

However, in addition to the specialization effect, an increase in the degree of intermediate inputs' differentiation also triggers other important effects. First, according to Kremer (1993), using more complex technologies (i.e., those that involve a larger number of production tasks and/or more varieties of an intermediate input) may hinder the manufacturing activity, the reason being that such technologies entail higher risks of failure. As a result, *complexity diseconomies* occur, as opposed to specialization economies. Second, the proliferation of varieties of intermediate inputs also implies a larger number of firms in the intermediate sector. Accordingly, the toughness of competition in the market for intermediate goods may increase or decrease, depending on the shape of demand for inputs. This effect, which we call *competition effect*, may also have a non-trivial impact on aggregate output and other key market variables.

In this paper, we undertake an in-depth inquiry into the nature of sharing externalities, and study how they affect market outcome. In more detail, we aim at showing that what is really key for understanding sharing externalities is the interplay between two forces: the specialization/complexity effect, on the one hand, and the competition effect, on the other hand. How the interaction between these two forces generates endogenous increasing returns to scale is definitely understudied in the literature, mainly because of the widely used assumption that technology in the final sector displays *constant elasticity of substitution* (henceforward CES) across the employed intermediates. Assuming CES technologies

is appealing as it simplifies enormously the description of market interactions between monopolistically competitive firms, thus increasing a model's tractability. The flipside of the coin, however, is that the equilibrium markup, which may serve as a reverse measure of the toughness of competition, remains unaffected by entry, as well as by market-size shocks. As a consequence, the competition effect is washed out. Both the "horizontal innovation" paradigm in endogenous growth theory (proposed first by Grossman and Helpman, 1990[1]), and the Marshallian externalities approach (first used by Abdel-Rahman and Fujita, 1990 to study agglomeration economies at the city level[2]) are essentially based on the CES assumption. For this reason, neither of these literatures allows to distinguish clearly between the impacts of specialization/complexity and toughness of competition on aggregate output and wages. The aim of our article is to fill this gap.

In order to achieve our purpose, we develop an extension of the Ethier's (1982) model to the case of a non-specified constant-returns-to-scale (CRS) technology in the final-good sector. The main result of our approach is a lucid decomposition of external increasing returns to scale into two components: (i) a *competition effect*, which stems from the market interactions between producers of intermediate inputs, and (ii) a *specialization/complexity effect*, which we model as a supply-side counterpart of the notion of "love/aversion for variety".[3] Through this decomposition, we show that the gains from specialization are, in general, *not the only* factor responsible for the emergence of external increasing returns to scale. Competition in the intermediate sector also plays a fundamental role, except when the final good is produced by means of a CES technology. In this common, yet very special, case external increasing returns to scale are *fully driven by the gains from specialization*. At the other extreme is the translog technology, where the external increasing returns to scale are induced *solely by the competition effect*. In between these two limit-situations, we find that both effects (specialization/complexity and competition) do matter in shaping market outcomes.

The other findings of the paper may be summarized as follows. First, we provide a micro-foundation of the complexity externality, which may lead to a reduction of TFP in the final output sector in response to expanding variety of intermediate inputs. Examples of how such an externality may work in the new growth theory can be found, just to mention some examples, in Howitt (1999), Dalgaard and Kreiner (2001), and Bucci (2013).

Second, we obtain a full characterization of the impact of horizontal innovation (which

[1]See also Krugman (1990, Ch. 11), Romer (1990), and Rivera-Batiz and Romer (1991).

[2]Duranton and Puga (2004) and Fujita and Thisse (2013) provide extensive and masterful surveys of this strand of literature.

[3]See Benassy (1996) and Zhelobodko et al. (2012) for recent models where "love for variety" plays a crucial role.

is embodied in the entry of new intermediate inputs' producers into the market) on prices, markups, and wages. Our model allows us to make a distinction between *price-decreasing/increasing* and *markup-decreasing/increasing* competition, and provides necessary and sufficient conditions for each type of competition to occur. We believe that this result represents a further theoretical advancement with respect to the treatment of the monopolistic competition process recently proposed by Zhelobodko et al. (2012). These authors, in fact, distinguish only between price-increasing and price-decreasing competition, the reason being that in their model prices and markups always move in the same direction in response to entry or exit of firms, as well as to market size shocks.

Third, we also discriminate between *wage-increasing* and *wage-decreasing* competition. Wage inequality is an issue which is important also from an international-trade perspective, as recently discussed, among others, by Amiti and Davis (2012) and Helpman et al. (2010). In this respect, our findings suggest that this kind of inequality may stem, at least in part, from the different nature of the interaction between the specialization/complexity and the competition effects across countries. Another issue that, in this regard, motivates empirically our theoretical analysis is the relationship between city size and wages. The exact form of this relationship is ambiguous, even though it is widely acknowledged by urban and regional economists that larger cities pay, on average, higher wages. In many papers, a log-linear relationship, implied by the CES model, is estimated. However, to improve the fit, city-specific dummies are commonly used (see, for example, Duranton, 2014). A more flexible strategy, for which our paper provides a microeconomic foundation, could be trying a non-linear specification.[4]

In addition, we find that the competition effect may either reinforce or weaken the impact of the specialization effect on aggregate output. This possibility is almost completely overlooked, for example, by horizontal R&D-based endogenous growth models (including Benassy, 1998) that consider only the positive effects of specialization, disregarding other possible effects (which may be both positive or negative in sign) which can come from an increase in the toughness of competition in the product market. In our analysis, the way in which the competition effect interacts with the specialization effect depends on whether the inverse demand elasticity for the intermediate inputs is a decreasing or an increasing function of the number of such inputs.

Finally, our main results are robust in the sense that they hold for any well-behaved technology which satisfies the (rather general) properties of symmetry and constant returns

[4]Needless to say, we acknowledge that factors other than specialization economies and market competition also play a significant role in determining the city size-wage gap. Moreover, Baum-Snow and Pavan (2012) clearly point out that this gap may be different across workers being heterogeneous in experience, ability, etc. However, these further dimensions of the problem are outside the scope of our paper.

to scale.

We are aware that extensions of our model along different lines and applications of it in many specific contexts are definitely possible. We briefly discuss some of these extensions and possible applications of our approach in the concluding section.

The article is structured as follows. In section 2 we set the model. In section 3 we characterize the equilibrium for a given number of input-producing firms. We also suggest a classification of competitive regimes in the intermediate input sector, based on the impact of entry on prices, markups, and wages. In section 4 we deal with a free-entry equilibrium, and study how the interaction between the specialization/complexity effect and the competition effect generates external increasing returns to scale. Section 5 concludes.

2 The Model

The economy is composed by two sectors that are involved into a vertical relationship. The intermediate inputs sector (henceforth sector $\mathcal{I}$), produces a differentiated intermediate good under monopolistic competition. The number of firms in this sector ($\mathcal{I}$-firms) is endogeneous due to free entry, while the only production factor is labor. Workers are homogeneous, and each of them inelastically supplies one unit of labor. The labor market is perfectly competitive.

The final good sector (hereafter sector $\mathcal{F}$) involves a unit mass of perfectly competitive firms ($\mathcal{F}$-firms) sharing the same CRS technology, which uses varieties of the intermediate good as inputs. The main departure of our modeling strategy from Ethier (1982) and other numerous subsequent papers lies in working with a non-specified production function instead of the widely used CES technology.

2.1 Sector $\mathcal{F}$

Production of the homogenous final good requires a continuum $[0, n]$ of varieties, henceforward *inputs*, of a horizontally differentiated intermediate good. All firms operating in sector $\mathcal{F}$ are endowed with the same production function F:

$$Y = F(\mathbf{q}), \tag{1}$$

where $\mathbf{q} = (q_i)_{i\in[0,n]}$ is the vector of intermediate inputs employed, while n stands for the number (more precisely, the *mass*) of intermediate inputs, as well as for the number of firms producing these inputs.

We make standard assumptions about $F(\mathbf{q})$. First, $F(\mathbf{q})$ is concave in $\mathbf{q}$, which implies that each input exhibits a diminishing marginal product. Second, $F(\mathbf{q})$ is positive homogenous of degree 1, so that returns to scale are constant. Finally, we focus on *symmetric* production functions, *i.e.* such that any permutation of intermediates does not change the final output, Y. The reason for imposing such a kind of symmetry, which typically holds in monopolistic competition contexts, is to refrain from placing any ad-hoc asymmetries on sector $\mathcal{I}$.

For illustrative purposes, we provide some examples.

1. CES: variations on a theme. The three assumptions just introduced (concavity, CRS, and symmetry) are simultaneously satisfied by the standard CES production function (Dixit and Stiglitz, 1977):

$$F(\mathbf{q}) \equiv \left(\int_0^n q_i^\rho \mathrm{d}i \right)^{1/\rho}, \quad 0 < \rho < 1. \tag{2}$$

One may think of at least two possible departures from production function (2). The first one is a production function where ρ is no longer a constant, but depends on n. This is the case studied by Gali (1995), who assumes that varieties become better and better technological substitutes as their number increases, i.e. $\rho'(n) > 0$. A second departure is the production function used in Ethier (1982):[5]

$$F(\mathbf{q}) = n^\nu \left(\int_0^n q_i^\rho \mathrm{d}i \right)^{1/\rho}, \quad 0 < \rho < 1. \tag{3}$$

In equation (3), when sufficiently negative, ν is a measure of the magnitude of the *complexity effect*: a larger number of intermediate inputs being simultaneously combined within the same production process can lead to a reduction in aggregate output (we come back to this issue immediately below). To be more precise, *complexity diseconomies* are said to occur if and only if $\nu < 1 - 1/\rho$. Otherwise ($\nu > 1 - 1/\rho$), *specialization economies* take place. The logic behind these definitions is as follows. Evaluating (3) at a symmetric output vector, in which $q_i = q$ for all $i \in [0, n]$, and where q is given, we obtain $Y = n^{\nu+1/\rho} q$. The above inequalitites keep track of whether Y increases more or less than proportionately with n. The baseline case described by (2) corresponds to $\nu = 0$, hence the CES technology can only account for specialization economies.

2. Translog production function. For a tractable example of a non-CES technology satisfying our assumptions, consider a production function given by

[5] See also Benassy (1998).

$$\ln F(\mathbf{q}) = \frac{1}{n}\int_0^n \ln q_i \mathrm{d}i - \frac{\alpha}{2n}\left[\int_0^n (\ln q_i)^2 \mathrm{d}i - \frac{1}{n}\left(\int_0^n \ln q_i \mathrm{d}i\right)^2\right] \tag{4}$$

One may treat (4) as an infinite-dimensional counterpart of the translog specification, which has been widely used in early empirical works on production functions estimation (see, e.g., Kim, 1992).

3. Kimball-type production functions. Kimball (1995) represents, to the best of our knowledge, one of the very first (and few) macroeconomic papers where a non-CES production technology is employed in sector $\mathcal{F}$. Namely, the production function $Y = F(\mathbf{q})$ is implicitly defined by means of the so-called "flexible aggregator":

$$\int_0^n \phi\left(\frac{q_i}{Y}\right) \mathrm{d}i = 1, \tag{5}$$

where $\phi(\cdot)$ is some function, which is assumed to be increasing and concave.[6]

Specialization economies vs complexity diseconomies. In order to extend the definitions of specialization economies and complexity diseconomies from the CES technology ($\nu \gtrless -1/\rho$) to any symmetric CRS technology, we consider the behavior of F at a symmetric outcome, i.e. when $q_i = q$ for all $i \in [0, n]$. Denote by $\varphi(n)$ the level of output that can be produced when a firm uses one unit of each intermediate input.[7]

Given total expenditure E allocated by a final good producer on the purchase of intermediate inputs under unit price for all of them, the *specialization economies* capture the idea that the division of labor generates productivity gains, namely a larger variety of intermediate inputs allows to produce a larger amount of final output. To put this in a more formal way, note that, due to constant returns to scale, output of the final good equals $q\varphi(n)$ when q units of each intermediate are employed. Hence, the specialization effect takes place if and only if

$$\frac{E}{n}\varphi(n) > \frac{E}{k}\varphi(k), \quad \text{where } k < n.$$

In other words, *specialization economies occur if and only if* $\varphi(n)/n$ *increases with* n, or equivalently when the elasticity of $\varphi(n)$ exceeds 1:

$$\frac{\varphi'(n)n}{\varphi(n)} > 1. \tag{6}$$

[6] To guarantee that a solution to (5) does exist for any n, one may assume additionally that $\phi(0) \leq 0$, while $\phi(\infty) = \infty$. When $\phi(\cdot)$ is a power function, we obtain the CES specification as a special case of (5).

[7] Formally, $\varphi(n) \equiv F\left(\mathbf{I}_{[0,n]}\right)$, where $\mathbf{I}_S$ is an indicator of $S \subseteq [0, n]$.

Otherwise output of the final good decreases with the intermediate inputs' range. In the latest case, we face *complexity diseconomies*.

In order to provide some intuition about how the trade-off between specialization economies/complexity diseconomies can take place, consider the following examples.

For the standard CES technology (2) we have $\varphi(n)/n = n^{(1-\rho)/\rho}$. Since $0 < \rho < 1$, this implies that just specialization economies take place. More generally, this is true for any production function satisfying (5) with $\phi(0) = 0$ (see Appendix 2).

For an example of a well-behaved technology which always displays complexity diseconomies, consider again the translog production function. Evaluating (4) at a symmetric input vector, we find that $\varphi(n) = 1$ for all $n > 0$. As a consequence, (6) is violated, which means the presence of complexity diseconomies.

Is it possible that specialization economies change into complexity diseconomies as n becomes sufficiently large? For this to happen, $\varphi(n)/n$ must be non-monotone. Apparently, the Kimball's "flexible aggregator" seems to be flexible enough to capture this possibility. To show this, consider a Kimball-type production function with $\phi(q/Y) \equiv a(q/Y)^\rho - b$, where $a, b > 0$, $0 < \rho < 1$. Here, a can serve as a measure of overall TFP, while b shows the strength of the complexity externality. This yields the following modification of the "augmented CES" technology:

$$F(\mathbf{q}) = A(n) \left(\int_0^n q_i^\rho \mathrm{d}i \right)^{1/\rho}, \quad A(n) \equiv \left(\frac{a}{1+bn} \right)^{1/\rho}. \tag{7}$$

The interaction between specialization and complexity underlying (7) is summarized by

$$\frac{\varphi(n)}{n} = \frac{1}{n} \left(\frac{an}{1+bn} \right)^{1/\rho}. \tag{8}$$

As implied by (8), $\varphi(n)/n$ is bell-shaped, [8] i.e. *specialization economies always prevail over complexity diseconomies whenever the intermediate input is not "too much" differentiated, otherwise complexity diseconomies always prevail over specialization economies.*

Cost function and price index. Each $\mathcal{F}$-firm seeks to minimize production costs,

$$\min_{\mathbf{q}} \int_0^n p_i q_i \mathrm{d}i \qquad \text{s.t.} \qquad F(\mathbf{q}) \geq Y, \tag{9}$$

treating Y as given. The approach based on the cost function $\mathscr{C}(\mathbf{p}, Y)$, which is defined as

[8] Indeed, it is readily verified that the right-hand side of (8) increases in n for all $n < 1/(\rho b)$ and decreases otherwise.

the value function of the cost minimization problem (9), provides a description of technology dual to the one based on the production function.[9] Because of constant returns to scale, a well-defined price index for intermediate goods $P(\mathbf{p})$ exists, which satisfies

$$\mathscr{C}(\mathbf{p}, Y) = YP(\mathbf{p}). \tag{10}$$

In the CES case, the price index is given by

$$P(\mathbf{p}) = \left(\int_0^n p_i^{1-\sigma} \mathrm{d}i \right)^{1/(1-\sigma)}. \tag{11}$$

An example of a price index which describes a tractable non-CES production function is the *translog price index*:

$$\ln P(\mathbf{p}) = \frac{1}{n} \int_0^n \ln p_i \mathrm{d}i - \frac{\beta}{2} \left[\int_0^n (\ln p_i)^2 \mathrm{d}i - \frac{1}{n} \left(\int_0^n \ln p_i \mathrm{d}i \right)^2 \right]. \tag{12}$$

Demand for inputs. The first-order condition for cost minimization is given by

$$p_i = \lambda \Phi(q_i, \mathbf{q}), \tag{13}$$

where $\Phi(q_i, \mathbf{q}) \equiv \partial F / \partial q_i$ is the marginal product[10] of input i, while λ is the Lagrange multiplier of the firm's program (9). It follows from the envelope theorem that the value of λ equals the marginal production cost, i.e. $\lambda = \partial \mathscr{C} / \partial Y$ for all Y and $\mathbf{p}$. Combining this with (10), we obtain the following inverse demand schedule for input i:

$$\frac{p_i}{P(\mathbf{p})} = \Phi(q_i, \mathbf{q}). \tag{14}$$

Weak interactions. As stated in the introduction, market interactions between producers of inputs are crucial for our results. For a better understanding of the nature of these interactions, a further inquiry on the properties of the marginal products $\Phi(q_i, \mathbf{q})$ is needed.

First, $\Phi(q_i, \mathbf{q})$ decreases in q_i, which is a straightforward implication of diminishing marginal returns. This property means that inverse demands (14) are downward-sloping.

[9] Duality theory in production, for the case of a finite set of inputs, was developed in pioneering works by Shephard (1953) and Uzawa (1964).

[10] A purist would correctly note at this stage that the partial derivatives $\partial F / \partial q_i$ are not well-defined in the case of a continuum of inputs, which may create some mathematical troubles in a framework where F is non-specified. We are aware of this problem. However, if we put some minor additional structure on the space of input vectors $\mathbf{q}$ potentially available for the final-good producers, everything works as if the marginal products were well-defined. See Appendix 1 for technical details.

Second, $\Phi(q_i, \mathbf{q})$ does not vary with individual output q_j of any firm $j \neq i$, given that the outputs of firm i and all the other firms (except j) remain unchanged (see Appendix 1 for details). This second property has a far-fetched implication: input-producing firms are not involved into truly strategic market interactions, but rather into *weak interactions*, meaning that the individual impact of each firm on the demand schedules of its competitors is negligible.[11] In other words, it is the aggregate behavior of firms to determine the market outcome, as no single firm has per se enough market power to strategically manipulate the market. This is typical in existing monopolistic competition models and is in the line with Chamberlin's "large group" assumption.

For the sake of illustation, consider again the CES case. The marginal products are given by

$$\Phi(q_i, \mathbf{q}) = q_i^{\rho-1}\mathcal{A}(\mathbf{q}), \qquad \mathcal{A}(\mathbf{q}) \equiv \left(\int_0^n q_j^\rho \mathrm{d}j \right)^{(1-\rho)/\rho}. \tag{15}$$

As implied by (15), in the CES case Φ is downward-sloping in q_i, while the demand shifter $\mathcal{A}(\mathbf{q})$ is invariant to individual changes in q_i.

A dual description of specialization/complexity (dis)economies. We now come to developing a dual description of the trade-off between specialization and complexity, as these two concepts have been defined before. To do so, we observe that when the price schedule for the intermediate inputs is symmetric, i.e. when $p_i = p$ for all $i \in [0, n]$, then the final-good producer will purchase all inputs in equal volumes: $q = Y/\varphi(n)$. As a consequence, total cost equals $Ypn/\varphi(n)$, while the price index at a symmetric outcome boils down to

$$P = \frac{n}{\varphi(n)}\, p. \tag{16}$$

Combining (16) with our definition of specialization economies, we may conclude that *the price index decreases (increases) with the range of inputs if and only if specialization economies (complexity diseconomies) take place.*

This dual approach to specialization economies/complexity diseconomies allows to see that the translog technology (12) is, in a sense, a *borderline* case. Indeed, as implied by (12), under a symmetric price schedule we have $P = p$. Comparing this with (16), we conclude that $\varphi(n) = n$. Thus, *neither specialization economies nor complexity diseconomies occur*,

[11] See Combes et al. (2008, Ch. 3) for a thorough discussion on the nature of weak interactions in monopolistic competition models.

i.e. these two forces exactly balance each other.[12]

The next proposition summarizes the main properties of all the production functions mentioned as examples so far.

Proposition 1. *(i) The augmented CES (3), the translog production function (4), the translog price index (12), and the Kimball's flexible aggregator (5) specify technologies which are all different with respect to each other, except the CES that can be obtained as a special case of production functions satisfying (5); (ii) Kimball-type technologies (5) satisfying* $\phi(0) = 0$ *exhibit specialization economies rather than complexity diseconomies; (iii) The translog production function* (4) *generates complexity diseconomies, while under the production function characterized by the translog price index (12) neither specialization economies nor complexity diseconomies take place.*

Appendix 2 provides the proof of Proposition 1. This result reveals the flexibility of our approach, which encompasses a wide variety of technologies with a differentiated input, including most of those that have been considered by previous works in the literature. In particular, our framework is more general than the one proposed by Kimball (1995), the reason being that, as implied by (i), the latter does not include the augmented CES, nor the translog production function.

2.2 Sector $\mathcal{I}$

There is a continuum of intermediate input producers sharing the same technology, which exhibits increasing returns to scale. Firm i's labor requirement for producing output q_i is given by $f + cq_i$, where $f > 0$ is the fixed cost and $c > 0$ is the constant marginal production cost. Thus, the profit π_i of firm i is defined by $\pi_i \equiv (p_i - cw)q_i - f$, where w stands for the wage rate.

Firm i faces the inverse demand schedule (14) and seeks to maximize its profit. Formally, the profit-maximization program of firm i is given by

$$\max_{p_i, q_i} [(p_i - cw)\, q_i] \quad \text{s.t.} \quad p_i = P(\mathbf{p})\phi(q_i, \mathbf{q}), \tag{17}$$

where $P(\mathbf{p})$ is the price index, which now plays the role of a market aggregate.

In accordance with the idea of weak interactions, individual changes in firms' prices have a negligible impact on $P(\mathbf{p})$, which is illustrated by (11) for the special case of the CES technology in sector $\mathcal{F}$. In other words, each $\mathcal{I}$-firm takes the value of P as a given. Hence,

[12] For another example of production function where the aggregate effect of variety expansion is suppressed (because the *specialization* and *complexity* consequences of an increase in the number of available varieties of intermediate inputs have the same magnitude but opposite sign), see Chen and Chu (2010, Eq. 4, p, 250).

(17) may be restated as

$$\max_{q_i} \left[\left(P\Phi(q_i, \mathbf{q}) - cw \right) q_i \right]. \tag{18}$$

The first-order condition for firm i's program (18) is given by

$$\Phi(q_i, \mathbf{q}) + q_i \frac{\partial \Phi}{\partial q_i} = \frac{cw}{P}. \tag{19}$$

Observe that the left-hand side of (19) is positive homogenous of degree zero. This implies that the solution of (19) cannot be unique. Indeed, multiplying a solution of (19) by a constant yields another solution. The "proper" equilibrium is pinned down by the labor balance condition

$$c \int_0^n q_i \mathrm{d}i + fn = L, \tag{20}$$

which equates total labor supply to total labor demand.

To guarantee that equation (19) is compatible with profit-maximizing behavior by firms, the second-order condition must hold, which amounts to assuming that the real operating profit $[\Phi(q_i, \mathbf{q}) - cw/P]\, q_i$ of firm i is strictly quasi-concave in q_i for all $\mathbf{q}$. Moreover, in order to ensure that a continuum of asymmetric Nash equilibria in the firms' quantity-setting game does not arise, we introduce a stronger assumption:

(Assumption A) *The left-hand side of (19) is decreasing in q_i for any $\mathbf{q}$.*

Imposing (Assumption A) is equivalent to assuming that the operating profit of each firm is strictly concave in its output. This assumption holds for the CES and, more generally, for any production function of the type (5) such that

$$-\frac{\phi'''(q/Y)}{\phi''(q/Y)} \frac{q}{Y} < 2 \quad \text{for all} \quad q/Y > 0.$$

(Assumption A) rules out asymmetric equilibria[13] because (19) has a unique solution $q_i^*(\mathbf{q})$, which is the same for all firms $i \in [0, n]$.

[13] See Gorn et al. (2012) for a recent formal treatment of multiple asymmetric equilibria in monopolistic competition.

3 Equilibrium for a given number of $\mathcal{I}$-firms

3.1 Equilibrium in sector $\mathcal{F}$.

Because the final output is consumed only by workers, product market balance suggests that $Y = wL$. This is possible only if $P = 1$. Indeed, firms' profits are given by $(1-P)Y$. Hence, if $P < 1$, each firm would supply infinitely many units of Y. On the contrary, if $P > 1$, total supply of the final good is zero, for no firm is willing to start production under negative profits.

Combining $P = 1$ with (16) pins down the equilibrium price for the intermediate inputs at a symmetric market configuration:

$$p^*(n) = \frac{\varphi(n)}{n}. \tag{21}$$

The intuition behind (21) is as follows. If the price for inputs exceeds $\varphi(n)/n$, then the supply of final good, hence the demands for inputs, are equal to zero. Consequently, firms producing intermediate goods will reduce prices in order to attract at least some demand. If, on the contrary, prices are lower than $\varphi(n)/n$, the supply of Y will be infinitely large, and so will be the demands for inputs, which would lead to a swell of prices.

Equation (21) may seem puzzling, as it implies that market interactions in sector $\mathcal{I}$ are fully irrelevant in determining input prices.[14]As a matter of fact, on the one hand it is absolutely true that the game between inputs' producers depends crucially on the market structure, and so do the profit-maximizing prices when the number of firms in endogenous (see Section 4.1 below). On the other hand, however, input-producing firms accurately anticipate the equilibrium value of the price index, which is determined outside the game among firms. Namely, it is driven to $P = 1$ by (i) perfect competition in sector $\mathcal{F}$, and (ii) correctness of the intermediate firms' expectations. In other words, under the assumption that the number of input-producing firms is given, things work *as if these firms were price-takers, even though they are actually price makers*. This property is a distinctive feature of Ethier's framework (as well as ours) compared to models of monopolistic competition a lá Dixit-Stiglitz (1977), where the final good is differentiated.

It is also worth mentioning that, because the labor market is perfectly competitive, the $\mathcal{I}$-firms take the wage w as given. Thus, the role of the equilibrium wage in this context is to align profit-maximizing prices with (21) (see Section 3.3).

Another important implication of (21) is that *the inputs' price at a symmetric equilibrium*

[14]Moreover, observe that (21) is fully independent of our assumption that sector $\mathcal{I}$ is monopolistically competitive. This relationship, indeed, would hold under any market structure which allows for a symmetric equilibrium (e.g., under symmetric Cournot or Bertrand oligopoly).

increases (decreases) with the number of firms n in sector $\mathcal{I}$ when specialization economies (complexity diseconomies) take place (see Section 2.1).

3.2 Equilibrium in sector $\mathcal{I}$.

Combining (14) with $P = 1$, the first order condition for profit maximization (19) may be recast as

$$\frac{p_i - cw}{p_i} = \eta(q_i, \mathbf{q}), \tag{22}$$

where η is the *marginal product elasticity*:

$$\eta(q_i, \mathbf{q}) \equiv -\frac{\partial \Phi}{\partial q_i} \frac{q_i}{\Phi(q_i, \mathbf{q})}. \tag{23}$$

At a symmetric outcome, when $p_i = p$ and $q_i = q$ for all $i \in [0, n]$, (22) boils down to

$$\frac{p - cw}{p} = r(n), \tag{24}$$

where $r(n)$ is defined by

$$r(n) \equiv \eta(q_i, \mathbf{q})|_{q_j = q_i \ \forall j \in [0,n]}. \tag{25}$$

Being a key ingredient of our model, $r(n)$ deserves some further comment. In general, $r(n)$ allows several different, though closely related, interpretations. First of all, $r(n)$ represents the *profit-maximizing markup*, which may serve as an inverse measure of the degree of product market competition. Thus, the behavior of $r(n)$ with respect to n shows how the toughness of competition varies with firm-entry. In particular, $r'(n) < 0$ would mean that competition gets tougher when more firms enter the market, which is probably the most plausible case, though not the only possible one. Second, as stated by (25), $r(n)$ is also the marginal product elasticity. In other words, $r(n)$ keeps track of whether the marginal product decreases at a higher or lower rate when the intermediate good becomes more differentiated. Finally, $r(n)$ also reflects the degree of product differentiation. Indeed, note that in the CES case we have

$$\frac{1}{r(n)} = \frac{1}{1 - \rho} = \sigma,$$

where σ is the *elasticity of technological substitution* across inputs. Hence, the higher σ, the lower product differentiation. In the non-CES case, setting $\sigma(n) \equiv 1/r(n)$ yields a measure of product differentiation which varies with n. It can be shown that $\sigma(n)$ is, in fact, the true elasticity of technological substitution across inputs (Nadiri, 1982) evaluated at a symmetric outcome.

In what follows, we prefer to view $r(n)$ mainly as the profit-maximizing markup. This interpretation is directly related to the notion of toughness of competition, which is of paramount importance for our results.

Before proceeding, it is legitimate to ask why $r(n)$ is independent of q. The key feature of the model which drives this result is the presence of constant returns to scale in sector $\mathcal{F}$. Indeed, due to the CRS property, $\Phi(q_i, \mathbf{q})$ is positive homogenous of degree zero (in Appendix 1, we provide a formal proof of this statement under a continuum of varieties). Combining this with (23), we find that η is also homogeneous of degree zero. This, in turn, implies that changing q in (25) under any given n does not perturb the right-hand side of (25). As a result, the profit-maximizing markup $r(n)$ depends solely on the number of firms.

The pricing rule (24) implies that competition gets tougher (softer) in response to entry of new $\mathcal{I}$-firms when $r(n)$ decreases (increases). In other words, competition may be either *markup-increasing* or *markup-decreasing*. Which of these two types of competition occurs in sector $\mathcal{I}$ is fully determined by the demand for inputs, which stems from sector $\mathcal{F}$. To illustrate this point, consider some examples. For the standard CES technology we have $r(n) = 1 - \rho$, i.e. profit-maximizing markups are not affected at all by entry of new $\mathcal{I}$-firms. Furthermore, for the cases of the translog production function (4) and the translog price index (12), the profit-maximizing markups are as follows:

Translog production function	Translog expenditure function
$r(n) = 1 - \alpha n$	$r(n) = \frac{1}{1+\beta n}$

Hence, both these technologies induce markup-decreasing competition.

When the production function is given by (5), we have

$$r(n) = -\frac{\xi \phi''(\xi)}{\phi'(\xi)}\bigg|_{\xi=\phi^{-1}(1/n)}, \tag{26}$$

i.e. competition is markup-decreasing if and only if the elasticity of $\phi'(\cdot)$ is an increasing function.

Specialization/complexity and competition. We now come to determining the equilibrium wages and aggregate final output, along with the equilibrium output per-firm. Combining (24) with (21) yields

$$w^*(n) = \frac{1}{c}[1 - r(n)]\frac{\varphi(n)}{n}. \tag{27}$$

Furthermore, plugging (27) into the product market balance $Y = Lw$, we obtain:

$$Y^*(n) = \frac{L}{c}[1 - r(n)]\frac{\varphi(n)}{n}, \tag{28}$$

Equations (27) and (28) are important because they suggest a decomposition of equilibrium wages and aggregate final output (up to the coefficients $1/c$ and L/c, respectively) into the product of the *competition effect*, $1 - r(n)$, and the *specialization/complexity effect*, $\varphi(n)/n$. The former increases with n if and only if $r'(n) < 0$, while the latter increases if specialization economies prevail over complexity diseconomies.

Finally, the per-firm output $q^*(n)$ is determined from the labor balance condition (20), which takes the form

$$(cq + f)n = L \tag{29}$$

at a symmetric outcome. Clearly, $q^*(n) = (L - fn)/(cn)$ always decreases with n.

3.3 The impact of entry on prices, wages, and markups

In our model, prices, wages, and markups are all endogenous. Moreover, putting together (21), (24), and (27), we observe that entry of new firms affects these variables not necessarily in the same direction. To be more precise, in what follows we say that competition is (i) *Price-decreasing* if $\partial p^*/\partial n < 0$, and *price-increasing* otherwise; (ii) *Markup-decreasing* if $\partial[(p^* - cw^*)/p^*]/\partial n < 0$, and *markup-increasing* otherwise; (iii) *Wage-decreasing* if $\partial w^*/\partial n < 0$, and *wage-increasing* otherwise.

The next Proposition summarizes the main results of the previous sub-section in terms of the above taxonomies.

Proposition 2. *In the framework of the model presented, competition is* (*i*) *Price-increasing* (*price-decreasing*) *if and only if the* $\mathcal{F}$*-firms enjoy specialization economies* (*suffer from complexity diseconomies*)*;* (*ii*) *Markup-decreasing* (*markup-increasing*) *if and only if* $r'(n) < 0$ ($r'(n) > 0$)*; and* (*iii*) *Wage-increasing* (*wage-decreasing*) *if and only if the following inequality holds* (*does not hold*):

$$\frac{\varphi'(n)n}{\varphi(n)} > 1 + \frac{r'(n)n}{1 - r(n)}. \tag{30}$$

Observe that, as implied by (27), the condition (30) is virtually $\mathrm{d}w^*/\mathrm{d}n > 0$ when written in terms of elasticities.

Proposition 2 highlights a crucial difference between our results and those recently obtained by Zhelobodko et al. (2012). These authors find that additional entry of firms leads to a reduction or hike in markups depending on how the elasticity of substitution varies with the individual consumption level. However, in their setting markup-decreasing competition is also price-decreasing and (because labor is chosen to be the numeraire) wage-decreasing,

and vice versa. In our model, this is not necessarily the case. To show this, we find it worth contrasting in a visual way our results about the impact of $\mathcal{I}$-firms' entry on prices, markups and wages across different types of production functions. Table 1 provides a summary for the CES and both types of translog technologies.

	Translog cost function	CES production function	Translog production function
Price	No effect	$\uparrow$	$\downarrow$
Markup	$\downarrow$	No effect	$\downarrow$
Wage	$\uparrow$	$\uparrow$	No effect

Table 1: The impact of entry on prices, markups and wages for different types of production functions

Table 1 reveals that under translog cost function prices are neutral to entry, while markups (wages) decrease (increase) in response to a larger number of firms. In the CES case, both prices and wages increase in response to more firms entering the intermediate input market, while the markup remains unchanged. Finally, with a translog production function wages remain unchanged when new firms enter, while both prices and markups fall. These findings highlight the key role of the interaction between the specialization/complexity effect and the competition effect in determining the nature of market outcomes.prices.[15]

4 External increasing returns to scale

This section describes in detail the interaction between the specialization/complexity effect and the competition effect in creating production externalities. Hence, it plays a central role within the whole analysis.

4.1 Free-entry equilibrium

We define a *symmetric free-entry equilibrium* as a vector $(p^*, q^*, n^*, w^*, Y^*)$, which satisfies (21), (24), (27), the labor balance condition (29), and the zero profit condition

[15] As an example of how our theory may be used, notice that both in trade and urban economics empirical evidence (???) tends to suggest that *larger cities exhibit higher prices, lower markups, and higher wages.* Table 1 reveals that neither the CES production function, nor any of the two translog technologies can fully capture this pattern. Proposition 2, however, suggests a qualified answer to the question of which production function would ultimately be able to reproducing these facts. Indeed, according to (30), if competition is both price-increasing and markup-decreasing, then it is also wage-increasing. Hence, any production function that exhibits *both* specialization economies ($n\varphi'(n)/\varphi(n) > 1$) and decreasing marginal product elasticities ($r'(n) < 0$) can replicate the empirical evidence for large cities mentioned above . This holds, in particular, for all Kimball-type production functions such that (i) $\phi(0) = 0$, and (ii) the elasticity of $\phi'(\cdot)$ is increasing.

$$(p - cw)q = wf. \tag{31}$$

Equilibrium number of firms. We first pin down the equilibrium number n^* of $\mathcal{I}$-firms. To do so, observe that (31) may be restated as follows:

$$\frac{p - cw}{p} = \frac{f}{f + cq}. \tag{32}$$

In other words, at a symmetric free-entry outcome the markup of any intermediate firm equals the share of fixed cost in firm's total production cost. This makes sense because it is the presence of a fixed cost which generates increasing returns to scale.

Combining (32) with the pricing rule (24) and the labor balance (29), we obtain:

$$r(n) = \frac{f}{L}n. \tag{33}$$

The equilibrium number of firms n^* is uniquely pinned down by (33) when $r(n)$ is either decreasing or "sufficiently slowly" increasing in n,[16] i.e. when competition is either markup-decreasing or "not too markup-increasing". If this is not the case, then multiple equilibria may arise. However, since we assume $0 < r(n) < 1$, (33) has always at least one solution $n^* > 0$, which means that a symmetric free entry equilibrium does always exist. In order to choose meaningful equilibria when they are multiple, we can restrict ourselves to *stable* equilibria, for which $r'(n^*) < f/L$.

Specialization and competition under free entry. Given n^*, using (24) and (31) yields the equilibrium firm's size:

$$q^* = \frac{f}{c}\frac{1 - r(n^*)}{r(n^*)} \tag{34}$$

According to (34), any shock that generates additional entry in the intermediate sector (and preserves the ratio f/c) would lead to a hike (respectively, a reduction) in firms' size if and only if $r(n)$ is a decreasing (respectively, increasing) function of n.

After plugging (34) into the production function of sector $\mathcal{F}$, we obtain the resulting *aggregate production function*:

$$Y^*(L) = \frac{L}{c}\left[1 - r\left(n^*(L)\right)\right]\frac{\varphi[n^*(L)]}{n^*(L)}, \tag{35}$$

while plugging n^* into (27) pins down the equilibrium wage w^*:

[16] By "sufficiently slowly" we mean that the elasticity of $r(n)$ never exceeds 1.

$$w^* = \frac{1}{c}\left[1 - r\left(n^*(L)\right)\right]\frac{\varphi\left[n^*(L)\right]}{n^*(L)}. \tag{36}$$

In equations (35) and (36), $1 - r\left(n^*(L)\right)$ captures the competition effect, which stems from sector $\mathcal{I}$. It increases (decreases) with n (and hence with the total labor supply, L) if and only if $r(n)$ is a decreasing (increasing) function of n, i.e. when competition is markup-decreasing (markup-increasing). The term $\varphi(n^*(L))/n^*(L)$ describes the specialization/complexity effect.

In order to clarify how the degree of competition in sector $\mathcal{I}$ may impact the sector $\mathcal{F}$'s aggregate production function, we observe that total output $Q^* \equiv n^* q^*$ in sector $\mathcal{I}$ is given by

$$Q\left[L,\, n^*(L)\right] = \frac{L}{c}\left[1 - f\,\frac{n^*(L)}{L}\right] = \frac{L}{c}\left[1 - r\left(n^*(L)\right)\right]. \tag{37}$$

Equation (37) follows from (29), (33), and (34). Using (37), the aggregate production function (35) may be restated as follows:

$$Y^*(L) = \frac{\varphi\left[n^*(L)\right]}{n^*(L)}\, Q\left[L,\, n^*(L)\right]. \tag{38}$$

The first term in (38) captures the specialization/complexity effect in sector $\mathcal{F}$, while the second term keeps track of the competition effect. In other words, in our framework *competition among input-producing firms affects total output of the final good through the aggregate output of the intermediate good.* More precisely, equations (33) and (37) imply that $Q\left[L,\, n^*(L)\right]$ increases more (respectively, less) than proportionally with L if and only if competition is markup-decreasing (respectively, markup-increasing). This, in turn, leads to competition generating a tendency toward external increasing (decreasing) returns to scale in sector $\mathcal{F}$. Compared to the standard CES model (where Q is readily verified to be exactly proportional to L, so that a competition effect cannot be taken into account), in the general case that we are analyzing *there are two sources of sharing externalities: the specialization/complexity effect* ***and*** *the competition effect.*

This explains why appealing to a CES production function may cause some limitations in various economic contexts. To see this in more detail, consider again equations (35) and (36) above. These two equations are basically the same and differ just by a constant term (L/c and $1/c$, respectively). When specialization economies take place, the term $\varphi[n^*(L)]/n^*(L)$ increases with n^*. As for the term capturing the competition effect, $[1 - r(n^*(L))]$, it rises with n^* under markup-decreasing competition ($r'(n^*) < 0$), and falls otherwise. Therefore, solely in the former case (markup-decreasing competition) the specialization effect on

both aggregate output and wages is reinforced by the competition effect. This is no longer true when competition is markup-increasing (in this case, the specialization effect would be weakened by the competition effect stemming from a larger number of firms entering the intermediate sector). Notice that, if the production function were CES, then the term $[1-r(n^*(L))]$, appearing in both (35) and (36), would be constant. Hence, the specialization effect would represent the only source of external increasing returns to scale in the final good sector.

4.2 The aggregate production function

We are now equipped to characterize the main properties of the aggregate production function, as well as to perform comparative statics of the free-entry equilibrium with respect to the number L of workers. Our main interest in this excercise is to reveal how aggregate output varies with L, namely how the external increasing/decreasing returns to scale in the $\mathcal{F}$-sector endogenously emerge. Under a positive shock in L, the left-hand side of equation (33) remains unchanged, while the right-hand side is shifted downwards. As a consequence, the equilibrium number of firms n^* increases with L whenever the equilibrium is stable, i.e. when $r'(n^*) < f/L$ (see Section 4.1). Combining this with (35), we find that at equilibrium the average product of labor, $Y^*(L)/L$, increases with L if and only if $[1 - r(n)]\varphi(n)/n$ is an increasing function of n, or, equivalently, if and only if competition is wage-increasing. Thus, we can now state the following proposition.

Proposition 3. *Endogenous increasing returns to scale take place if and only if (30) holds, or, equivalently, competition is wage-increasing.*

As discussed in Section 3, what renders competition wage-increasing or wage-decreasing in our model is the *interplay* between the competition effect and the specialization/complexity effect. Therefore, Proposition 3 stresses the importance of the interaction between the two effects in generating Marshallian externalities. Indeed, by comparing (30) with (6), which is a necessary and sufficient condition for specialization economies to arise, one immediately notices that the major difference between the two conditions resides in the fact that the former contains an additional term, $nr'(n)/[1-r(n)]$, which is due to the competition effect. In particular, it is possible to see that (30) and (6) coincide if and only if $r(n)$ is constant, which corresponds to the classical case of a CES technology. This explains why both endogenous growth and agglomeration economics literatures have generally explained so far the emergence of external increasing returns to scale by appealing solely to the presence of specialization economies, so almost totally (and perhaps undeservedly) neglecting the role of market interactions among firms in this process.

Concerning the relationship between n^* and L, our analysis reveals that n^*increases less (more) than proportionally in L if and only if $r(n)$ is a decreasing (increasing) function of n. In other words, when competition is markup-decreasing (markup-increasing), the number of firms increases less (more) than proportionally in response to an increase in L. Combining this with (35), we can state the following result.

Proposition 4. *Compared to the CES case, markup-decreasing competition damps the specialization effect, but simultaneously triggers a positive competition effect. Under markup-increasing competition, the situation is reversed.*

Table 2 summarizes in a compact way our results about the roles that market-size and the interaction between the specialization effect and the competition effect play in determining the equilibrium market-outcome under markup-decreasing and markup-increasing competition:

	$r'(n) < 0$	$r'(n) > 0$
n^*	increases less than proportionally in response to an increase in L	increases more than proportionally in response to an increase in L
$Y^*,\ w^*$	specialization effect weakened, positive competition effect	specialization effect reinforced, negative competition effect

Table 2: The impact of market-size and the interplay between the competition and the specialization effects in determining the equilibrium market-outcome, depending on the nature of competition (*markup-decreasing* vs. *markup-increasing*)

To illustrate further the role of the interaction between the specialization/complexity effect and the toughness of competition in shaping the aggregate production function, consider some examples.

CES production function. In this case, equation (33) is linear, i.e. the number of firms is proportional to total labor supply L. Hence, the competition effect is washed out, and the specialization effect is the only source of external increasing returns. The aggregate production function is given by

$$Y^*(L) = AL^{1/(1-\rho)}, \quad A \equiv \frac{\rho}{c}\left[\frac{1-\rho}{f}\right]^{\rho/(1-\rho)}.$$

Translog price index. Combining (12) with (33) yields $\beta n^2 + n = L/f$, which implies $n^* = \left(\sqrt{1+4L/f} - 1\right)/(2\beta)$. In this case the number of firms grows proportionally to $\sqrt{L}$. This is because, unlike the CES case, here competition becomes tougher and tougher as the market gets larger. Hence, following an increase in the number of firms, complexity diseconomies and specialization economies exactly offset each other. Meanwhile, the competition

effect becomes the main force shaping the resulting aggregate production function which is given by

$$Y^*(L) = \frac{f}{4\beta c}\left(\sqrt{1+4L/f}-1\right)^2. \tag{39}$$

Equation (39) suggests that the average product of labor $Y^*(L)/L$ increases in L for all $L \geq 0$. In other words, external increasing returns take place. However, the source of these increasing returns is radically different from the one in the CES case. Namely, *agglomeration economies stem here solely from market interactions between firms,* while in the classical CES-based models they are generated entirely by technological externalities embodied in the specialization/complexity tradeoff.

Translog production function. In this case, the competition effect is even stronger. Indeed, as implied by (4), (33) takes the form: $1-\alpha n = fn/L$. Hence, $n^* = L/(\alpha L + f)$, which implies that the equilibrium number of firms is bounded from above by $1/\alpha$. In other words, even when L grows unboundedly, the number of firms the market invites to operate remains limited due to very tough competition. The aggregate production function is given by

$$Y^*(L) = \frac{\alpha}{c}L. \tag{40}$$

Thus, in the case of translog production function, the resulting technology exhibits *constant returns to scale.* This result stems from the fact that, according to Proposition 3, agglomeration economies/diseconomies arise only when competition is wage-increasing/decreasing. As found in Section 3.2 (Table 1), under translog production function entry has no impact on wages.

A micro-foundation for an S-shaped aggregate production function. Consider again the production function given by (7). The resulting aggregate production function $Y^*(L)$ reads as

$$Y^*(L) = \frac{f}{1-\rho}\left(\frac{L}{L+af/(1-\rho)}\right)^{1/\rho}. \tag{41}$$

According to (41), increasing returns to scale arise when L is sufficiently small; otherwise, decreasing returns to scale occur. Thus, (7) provides a simple micro-foundation for an S-shaped aggregate production function, which dates back to Shapley and Shubik (1967) and has been widely used in growth theory and development economics, especially in the analysis of poverty traps.[17]

[17]See Skiba (1978), and, more recently, Azariadis and Stachurski (2005), as well as Banerjee and Duflo (2005), for examples on the possible consequences of using S-shaped production functions within these two

5 Concluding remarks

The *"CES-paradigm"* is now among the best-established ones in micro- and macro-economic theory. However, while easy to handle, this paradigm fails to capture the fact that the substitutability across varieties of an intermediate input may not be constant, but linked to the evolution of the number of these varieties/the number of firms producing these inputs. Using a two-sector model in which a perfectly competitive final good sector and a monopolistically competitive intermediate input sector are vertically integrated, we have disentangled the interplay between two effects: the specialization/complexity effect, arising from the final output sector (employing more varieties of intermediate inputs fosters/deters production of the final good), and the competition effect, stemming from the market interactions among firms within the intermediate input sector. The latter effect is driven by the presence of a variable elasticity of technological substitution and would vanish in the limiting case in which technology were CES.

In our model, the nature of product market competition and the interaction between the competition effect and the specialization/complexity effect are determined by the behavior of the marginal product elasticity. More specifically, depending on whether the marginal product elasticity decreases or increases with the number of inputs at a symmetric outcome, competition may be either markup-increasing or markup-decreasing. This distinction is decisive in our framework because, conditional on the type of competition, we find that: (i) The competition effect can be either positive or negative, thus it can either strengthen or weaken the specialization effect on both equilibrium total output and wages; (ii) The equilibrium number of firms can increase more or less than proportionally in response to a rise in the total labor supply, which is generally regarded as a proxy for the market size; (iii) Whether external increasing returns to scale do emerge or not depends crucially on the interaction of the competition effect with the specialization/complexity externality, not just on the degree of gains from specialization.

We believe that these results may be useful for further advancements in economic models based on the joint presence of increasing returns to scale and imperfect competition in the product market, which are extensively used in international trade and agglomeration, economic geography, and endogenous growth. Furthermore, our treatment of production functions may also turn out to be helpful in development economics and the analysis of poverty traps. In order to better fit our approach to current research agenda in these various fields of economics, several extensions of the baseline model come to mind. To mention just a few, one may think of (i) accounting for heterogeneity of firms in productivity, and

branches of economic literature.

studying the role of selection among firms in generating sharing externalities, (ii) more realistic frameworks in which more than two productive sectors are simultaneously active in the economy, and (iii) considering market structures different from monopolistic competition in the input-producing sector. We leave the formal development of these more specialized issues for future research.

References

[1] Abdel-Rahman, H., and M. Fujita (1990). Product variety, Marshallian externalities, and city sizes. *Journal of Regional Science* 30: 165-183.

[2] Amiti, M., and D.R. Davis (2012). Trade, firms, and wages: Theory and evidence. *Review of Economic Studies* 79: 1-36.

[3] Azariadis, C, and J. Stachurski (2005). "Poverty Traps". In P. Aghion and S.N. Durlauf (Eds.), Handbook of Economic Growth. Amsterdam: Elsevier-North Holland, Chap. 5 (Volume 1A): 295-384.

[4] Baum-Snow, N., and R. Pavan (2012). Understanding the city size wage gap. *Review of Economic Studies* 79: 88-127.

[5] Banerjee, A.V., and .E. Duflo (2005). "Growth Theory through the Lens of Development Economics". In P. Aghion and S.N. Durlauf (Eds.), Handbook of Economic Growth. Amsterdam: Elsevier-North Holland, Chap.7 (Volume 1A): 473-552.

[6] Benassy, J.-P. (1996). Taste for variety and optimum production patterns in monopolistic competition. *Economics Letters* 52: 41-47.

[7] Benassy, J.-P. (1998). Is there always too little research in endogenous growth with expanding product variety? *European Economic Review* 42: 61-69.

[8] Bucci, A. (2013). Returns to specialization, competition, population, and growth. *Journal of Economic Dynamics & Control* 37: 2023-2040.

[9] Chen, B.-L., and A.C. Chu (2010). On R&D spillovers, multiple equilibria and indeterminacy. *Journal of Economics* 100: 247-263.

[10] Combes, P.P., Mayer, T., and J.F. Thisse (2008). *Economic geography: The integration of regions and nations.* Princeton University Press.

[11] Dalgaard, C.-J., and C.T. Kreiner (2001). Is declining productivity inevitable? *Journal of Economic Growth* 6: 187-203.

[12] Dixit, A.K., and V. Norman (1980). *The Theory of International Trade*. Cambridge, UK: Cambridge University Press.

[13] Dixit, A.K., and J.E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *American Economic Review* 67: 297-308.

[14] Duranton, G., and D. Puga (2004). *Micro-foundations of urban agglomeration economies*. Handbook of Regional and Urban Economics, Ch. 4: 2063-2117.

[15] Duranton, G. (2014). Growing through cities in developing countries. *The World Bank Research Observer* (first published online: April 15, 2014, doi:10.1093/wbro/lku006).

[16] Ethier, W. J. (1982). National and international returns to scale in the modern theory of international trade. *The American Economic Review* 72: 389-405.

[17] Fujita, M., and J.F. Thisse (2013). *Economics of Agglomeration: Cities, Industrial Location, and Globalization. Second Edition*. Cambridge, UK: Cambridge University Press.

[18] Gali, J. (1995). Product diversity, endogenous markups, and development traps. *Journal of Monetary Economics* 36: 39-63.

[19] Gorn, A., Kokovin, S., and E. Zhelobodko (2012). Non-concave Profit, multiple asymmetric equilibria and catastrophes in monopolistic competition, *mimeo*.

[20] Grossman, G.M., and E. Helpman (1990). Comparative advantage and long run growth. *American Economic Review* 80: 796–815.

[21] Helpman, E., O. Itskhoki, and S. Redding (2010). Inequality and unemployment in a global economy. *Econometrica* 78: 1239-1283.

[22] Howitt, P. (1999). Steady endogenous growth with population and R&D iputs growing. *Journal of Political Economy* 107: 715–730.

[23] Kim, H. Y. (1992). The translog production function and variable returns to scale. *Review of Economics and Statistics*: 546-552.Kimball, M.S. (1995). The quantitative analytics of the basic nonmonetarist model. *Journal of Money, Credit and Banking* 27: 1241–1277.

[24] Kremer, M. (1993). The O-Ring Theory of Economic Development. *Quarterly Journal of Economics* 108: 551-575.

[25] Krugman, P.R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics* 9: 469–479.

[26] Krugman, P.R. (1981). Intraindustry specialization and the gains from trade. *Journal of Political Economy* 89: 959–973.

[27] Krugman, P.R. (1990). *Rethinking International Trade*. Cambridge, MA: MIT Press.

[28] Lancaster, K. (1980). Intraindustry trade under perfect monopolistic competition. *Journal of International Economics* 10: 151–175.

Nadiri, M.I. (1982) Producers theory. In Arrow, K.J. and M.D. Intriligator (eds.) *Handbook of Mathematical Economics. Volume* II. Amsterdam: North-Holland, pp. 431 – 90.

[29] Rivera-Batiz, L.A. and P.M. Romer (1991). Economic integration and endogenous growth. *Quarterly Journal of Economics* 106: 531-555.

[30] Romer, P.M. (1990). Endogenous technological change. *Journal of Political Economy* 98: S71–S102.

[31] Shapley, L. S., and M. Shubik (1967). Ownership and the production function. *The Quarterly Journal of Economics*: 88-111.

[32] Shephard, R. W. (1953). *Cost and production functions*. Princeton University Press, NJ.

[33] Skiba, A.K. (1978). Optimal growth with a convex-concave production function. *Econometrica* 46: 527-539.

[34] Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London: W. Strahan and T. Cadell.

[35] Uzawa, H. (1964). Duality principles in the theory of cost and production. International Economic Review, 5(2), 216-220.

[36] Zhelobodko, E., S. Kokovin, M. Parenti, and J.F. Thisse (2012). Monopolistic competition: beyond the constant elasticity of substitution. *Econometrica* 80: 2765–2784.

Appendices

Appendix 1. Marginal products under a continuum of inputs

We restrict our attention to such input vectors $\mathbf{q}$ that have a finite second moment, i.e. $\int_0^n q_i^2 \mathrm{d}i < \infty$. In other words, $\mathbf{q} \in L_2([0,n])$. Intuitively, this assumption allows *mean* and *variance* of the input vector to be well-defined.

We also assume Frechet-differentiability, i.e. we postulate that there exists a functional $\Phi : \mathbb{R}_+ \times L_2 \to \mathbb{R}_+$, such that

$$F(\mathbf{q}+\mathbf{h}) = F(\mathbf{q}) + \int_0^n \Phi(q_i, \mathbf{q}) h_i di + \circ(||\mathbf{h}||_2) \text{ for all } \mathbf{q}, \mathbf{h} \in L_2. \tag{42}$$

In Eq. (42), $||\cdot||_2$ stands for the L_2-norm, i.e. $||\mathbf{h}||_2 \equiv \sqrt{\int_0^n h_i^2 \mathrm{d}i}$, whereas $\Phi(q_i, \mathbf{q})$ is the *marginal product* of intermediate input i. Concavity of F implies that Φ is decreasing in q_i.

Lemma. *Let $F : L_2 \to \mathbb{R}_+$ be a Frechet-differentiable functional, which is positive homogeneous of degree* 1. *Then* (i) *$\Phi(q_i, \mathbf{q})$ is positive homogenous of degree zero in $(q_i, \mathbf{q})$, and* (ii) *the Euler's identity*

$$F(\mathbf{q}) = \int_0^n q_i \Phi(q_i, \mathbf{q}) \mathrm{d}i, \tag{43}$$

holds.

Proof. To prove (i), rewrite (42) as follows:

$$F(t\mathbf{q}+t\mathbf{h}) = F(t\mathbf{q}) + \int_0^n \Phi(tq_i, t\mathbf{q}) t h_i di + \circ(t||\mathbf{h}||_2) \text{ for all } \mathbf{q}, \mathbf{h} \in L_2,\ t \in \mathbb{R}_+. \tag{44}$$

Dividing both sides of (44) by t and using homogeneity of F, we obtain

$$F(\mathbf{q}+\mathbf{h}) = F(\mathbf{q}) + \int_0^n \Phi(tq_i, t\mathbf{q}) h_i di + \circ(||\mathbf{h}||_2). \tag{45}$$

Combining (42) with (45), we find that $\phi(tq_i, t\mathbf{q})$ is a Frechet derivative of F computed at $\mathbf{q}$ for any $t > 0$. By uniqueness of Frechet derivative, $\phi(tq_i, t\mathbf{q})$ must be independent of t, which proves part (i) of the Lemma.

To prove part (ii), note that (42) implies the following identity:

$$\frac{F\left((t+\tau)\mathbf{q}\right) - F(t\mathbf{q})}{\tau} = \int_0^n \Phi(tq_i, t\mathbf{q}) q_i di + \frac{\circ(\tau)}{\tau} \text{ for all } \tau \in \mathbb{R}. \tag{46}$$

Using homogeneity of F and Φ, we obtain (43) as the limiting case of (46) under $\tau \to 0$. Q.E.D.

Appendix 2. Proof of Proposition 1.

We find it convenient to prove (i) – (iii) in reverse order.

As shown in Section 2.1, under (4), (respectively (12)), we have $\varphi(n) = 1$, (respectively, $\varphi(n) = n$) for all $n > 0$. Thus, claim (iii) follows immediately.

We now move to verifying claim (ii). If a production function satisfies (5), we have

$$\frac{\varphi(n)}{n} = \frac{1/n}{\phi^{-1}(1/n)}. \tag{47}$$

Because $\phi(\cdot)$ is increasing and concave, it must be that $\phi^{-1}(\cdot)$ is increasing and convex. If $\phi(0) = 0$, then the elasticity of $\phi^{-1}(\cdot)$ always exceeds 1. As a consequence, $\varphi(n)/n$ decreases in $1/n$ and increases with n.

When $\phi(0) \neq 0$, the above argument is no longer valid. Indeed, as implied by (8), production function given by (7) provides a counterexample. This completes the proof of (ii).

Finally, to prove claim (i), we proceed in four steps.

Step 1. Assume there exists an $\alpha > 0$ and an increasing convex function $\phi(\cdot)$, such that the translog production function (4) satisfies (5). When production is given by (4), then $\varphi(n)/n = 1/n$. Comparing this with (47) yields that $\phi^{-1}(\cdot)$ must be a constant, which contradicts the fact that it is strictly increasing.

Step 2. Assume that (4) belongs to (5). More precisely, let a $\beta > 0$ and a function ϕ do exist, such that the cost function dual to (5) is given by (12). Under (12), $\varphi(n)/n = 1$. Comparing this with (47) yields that $\phi^{-1}(\cdot)$ must be the identity mapping, and so is $\phi(\cdot)$. In this case, (5) defines a linear production function, for which the dual cost function is Leontief, not translog.

Step 3. Since (4) and (12) give rise to different profit maximizing markups, these two

technologies also differ with ech other.

Step 4. To show that (3) cannot satisfy (5) when $\nu \neq 0$, we first define a *generalized augmented CES* production function:

$$F(\mathbf{q}) \equiv A(n) \left(\int_0^n q_i^\rho \, di \right)^{1/\rho}, \quad 0 < \rho < 1, \tag{48}$$

where $A(n)$ is an exogenous externality. We now show that if a production function satisfies simultaneously (5) and (48) for some functions $A(\cdot)$ and $\phi(\cdot)$, then it must be that $A(n) = [a/(1+bn)]^{1/\rho}$, i.e. (48) coincides with (7). Indeed, the multiplier $A(n)$ does not have an impact on $\eta(q_i, \mathbf{q})$. Hence, $r(n) = 1 - \rho$. Combining this with (5), which implies

$$r(n) = -\frac{\phi''[\phi^{-1}(1/n)]}{\phi'[\phi^{-1}(1/n)]} \phi^{-1}(1/n),$$

we conclude that $\phi'(\cdot)$ must be a power function. This, in turn, implies $\phi(q/Y) \equiv a(q/Y)^\rho - b$. As a result, (7) holds. Because it is impossible under $\nu \neq 0$ that $[a/(1+bn)]^{1/\rho} = n^\nu$ for all n, this proves that (3) and (5) cannot be satisfied together when $\nu \neq 0$.

Setting $\nu = b = 0$ and $a = 1$, we obtain the standard CES production function as a special case of both (7) and (3). Otherwise, (3), (4), (12), and (5) specify different technologies. Q.E.D.