On the Microeconomics of Specialization: An Application to Agriculture

by

Jean-Paul Chavas

and

Kwansoo Kim[1]

Abstract: This paper investigates the microeconomics of specialization and its effects on firm productivity. We define economies of specialization as the productivity gains obtained under greater specialization. The paper shows how scale effects and non-convex technology affect economies of specialization. Using a nonparametric approach, we present an empirical analysis applied to Korean farms. The results indicate that non-convexity is prevalent especially on large farms. We find that non-convexity generates large productivity benefits from specialization on larger farms (but not on smaller farms), providing a strong incentive for large farms to specialize. We evaluate the linkages between non-convexity, firm size and management.

Keywords: firm productivity, scale, non-convexity, specialization

J.E.L.: D2, L25, Q12

[1] Jean-Paul Chavas is Professor, Taylor Hall, University of Wisconsin, Madison, WI 53706, USA; Email: jchavas@wisc.edu. Kwansoo Kim is Professor, Seoul National University, Seoul, Korea; Email: kimk@snu.ac.kr. This research was funded in part by a grant from the Graduate School, University of Wisconsin, Madison.

On the Microeconomics of Specialization: An Application to Agriculture

1. Introduction

Adam Smith (1776) first pointed out that there are productivity gains from specialization. Using a pin factory as an example, Smith (1776, p. 4) argued that producing pins in a system where workers are specialized across tasks can generate very large increases in productivity. According to Smith (1776, p. 6-8), a key factor is the amount of time workers spend switching from one task to another: this time can be saved under increased specialization. Other aspects of the benefits of firm specialization relate to the role of knowledge and coordination cost (e.g., as emphasized by Becker and Murphy (1992) and Caliendo and Rossi-Hansberg (2012)). Somewhat surprisingly, little empirical evidence has been presented documenting the source or magnitude of gains from specialization at the firm level. This suggests a need for a refined analysis of the productivity effects of specialization. The main objective of this paper is to develop new insights into the microeconomics of firm organization and the motivations for firm diversification/specialization strategies.[1]

Where do the productivity gains from firm specialization come from? This paper shows that there are two main factors that affect the benefits of specialization: returns to scale and nonconvexity of the technology. The role of returns to scale is not new: it has been noted in previous literature (e.g., Stigler 1951, Krugman, 1980, Melitz, 2003). But other factors also play a role. Smith's example of a pin factory provides useful insights. As noted above, Smith (1676, p. 6-8) argued that the gains from specialization come in part from saving in time lost switching from one task to the next. A similar argument would apply to the time used in learning how to manage a new task (Becker and Murphy, 1992; Caliendo and Rossi-Hansberg, 2012; Coviello et al., 2014). To the extent that the time lost switching between tasks (or learning to manage a new task) does not contribute to any output, this introduces fixed cost in the analysis. This suggests that reductions in fixed cost can be important sources of gains from specialization. This point has been made by Baumol et al. (1982, p. 75) in their analysis of economies of scope for a multiproduct firm. Fixed cost is a well-known source of non-convexity. Perhaps more importantly, these issues can arise independently of any scale effects. For example, in his discussion of a pin factory, Smith does not mention any role for firm size. This indicates that the productivity effects of specialization can be present within a firm irrespective of scale effects. This suggests a need to explore the role of non-convexity in the microeconomics of firm specialization.

This paper explores the microeconomics of specialization, with a focus on the role of non-convexity. It is well known that a technology that exhibits increasing returns to scale (IRS) is also non-convex. Yet, in this paper we stress that non-convexity can arise in ways that are unrelated to scale effects. Indeed, IRS is a form of non-convexity that applies in a very restrictive way: returns to scale consider only proportional changes in all inputs and outputs. We show that other forms of non-convexity (besides IRS) can have a large influence on the gains from specialization. More fundamentally, we think that the common idea that IRS and non-convexity tend to go together has contributed to hiding the deeper role played by non-convexity.

This paper makes three contributions to the literature. First, it evaluates conceptually the role played by both returns to scale and non-convexity in the economics of firm specialization. Relying on a directional distance function, we propose a measure of gains from specialization and use it to identify the distinct role played by returns to scale versus non-convexity. We obtain the following key result: the gains of firm specialization are negative under increasing returns to scale and a convex technology; alternatively, the gains of firm specialization are positive under decreasing returns to scale and a non-convex technology. Thus, our analysis shows that non-convexity can be an important factor contributing to the gains from firm specialization. This indicates a need to assess empirically the nature of returns and scale and non-convexity for a firm.

Our second contribution is to study the effects of non-convexity on specialization incentives. The analysis is based on a general measure of non-convexity. The measure is evaluated empirically using a non-parametric method. The non-parametric method is flexible in the sense that it allows for the presence of non-convexity in any part of the technology.[2] It provides a good basis to evaluate the role of management in firm specialization decisions.

Our third contribution is to apply our approach to a sample of Korean farms. An application to farms is of interest as most farms produce more than one output, allowing us to observed different patterns of output specialization across farms. In addition, farms are typically family farms where the head of the household is the manager and most labor is provided by family labor. In this case, we can expect the gains from specialization to be closely associated with the managerial skills of the farm manager, i.e. his ability to manage multiple farm production activities. Our empirical analysis documents the relative role played by returns to scale and non-convexity on Korean farms. The results identify the presence of non-convexity as well as scale effects. We show that non-convexity varies across farm types: non-convexity tends to be more common on larger farms. We also find that non-convexity effects are more important than scale effects on larger farms. It means that scale effects are not likely to be the major/single factor affecting firm specialization (as documented below). By showing how non-convexity varies with farm size, our analysis provides useful insights into the microeconomics of specialization. It helps explain why larger farms tend to be more specialized (Chavas, 2001). Finally, our application evaluates the linkages between management and non-convexity. We find that non-convexity varies with the education and experience of the farm manager. We also find that non-convexity generates large productivity benefits from specialization on larger farms (but not on smaller farms), providing a strong incentive for large farms to specialize. By evaluating the linkages between non-convexity, firm size and management, our analysis provides new insights into the role of management and the economics of specialization.

The paper is organized as follows. Section 2 presents a conceptual approach to the economics of firm specialization, with a focus on the role of returns to scale and non-convexity. Section 3 presents specific measures of economies of scale and non-convexity. And section 4 shows how these measures can be evaluated empirically using nonparametric methods. Section 5 presents an empirical application to a data set of Korean farms. Finally, section 6 concludes.

2. The Microeconomics of Specialization

Consider a production process involving m netputs $z \equiv (z_{1,} \ldots, z_m) \in R^m$. Given $z \equiv (z_{1,} \ldots, z_m)$, we use the netput notation where inputs are negative ($z_i \leq 0$ for input i) and outputs are positive ($z_j \geq 0$ for output j). The production technology is represented by the feasible set $T \subset R^m$, where $z \in T$ means that the netput vector z is feasible. The set T provides a global characterization of the underlying technology. Two specific properties of the technology will be examined in this paper: returns to scale and convexity properties. First, the technology T is said to exhibit $\left\{\begin{array}{c}\text{increasing returns to scale (IRS)}\\ \text{constant returns to scale (CRS)}\\ \text{decreasing returns to scale (DRS)}\end{array}\right\}$ if $T \left\{\begin{array}{c}\supset\\ =\\ \subset\end{array}\right\} \delta\, T$ for any scalar $\delta > 1$; and the technology is said to exhibit variable returns to scale (VRS) if no *a priori* restriction is imposed on returns to scale. Second, the technology is said to be convex if the set T is convex, i.e. if it satisfies $[\alpha\, z^a + (1-\alpha)\, z^b] \in T$ for any $z^a \in T$, $z^b \in T$, and $\alpha \in [0, 1]$. A convex technology is equivalent to the intuitive concept of "decreasing marginal productivity." Alternatively, the technology is non-convex if the set T is not convex. Throughout the paper, we assume that the technology T satisfies free disposal, where free disposal means that $T = T - R_+^m$.

Our analysis of the properties of the technology T will rely on specific measures. Letting $g \in R_+^m$ be a reference bundle satisfying $g \neq 0$ and following Chambers et al. (1996), consider the directional distance function[3]

$$D(z, T) = \sup_\beta \{\beta: (z + \beta\, g) \in T\} \text{ if there is a scalar } \beta \text{ satisfying } (z + \beta\, g) \in T\}, \quad (1)$$

$$= -\infty \text{ otherwise.}$$

The directional distance function is the distance between point z and the upper bound of the technology T, measured in number of units of the reference bundle g. It provides a general measure of productivity. In general, $D(z, T) = 0$ means that point z is on the frontier of the technology T. Alternatively, $D(z) > 0$ implies that z is technically inefficient (as it is below the frontier).[4] And $D(z, T) < 0$ identifies z as being infeasible (as it is located above the frontier). Luenberger (1995) and Chambers et al. (1996) provide a detailed analysis of the properties of D(z, T). First, by definition in (1), $z \in T$ implies that $D(z, T) \geq 0$ (since $\beta = 0$ would then be feasible in (1)), meaning that $T \subset \{z: D(z, T) \geq 0\}$. Second, $D(z, T) \geq 0$ in (1) implies that $[z + D(z, T)\, g] \in T$. When the technology T exhibits free disposal, it follows that $D(z, T) \geq 0$ implies that $z \in T$, meaning that $T \supset \{z: D(z, T) \geq 0\}$. Combining these two properties, we obtain the following result: under free disposal, $T = \{z: D(z, T) \geq 0\}$ and D(z, T) provides a complete representation of the technology T. Importantly, besides being convenient, this result is general: it allows for an arbitrary multi-input multi-output technology; and it applies with or without convexity.[5]

The distance function D(z, T) in (1) can be used to evaluate economies of specialization. To see that, consider two situations: one where netput z is produced by a single firm; and one where z is produced by K more specialized firms, where the k-th firm produced z^k subject to the restriction $\sum_{k=1}^{K} z^k = z$. Here, the constraint $\sum_{k=1}^{K} z^k = z$ requires that the aggregate netputs are the same in both situations.

Definition 1: At points z and $(z^1, ..., z^K)$ satisfying $z = \sum_{k=1}^{K} z^k$, define the following measure of economies of specialization

$$EP(z, z^1, ..., z^K, T) = \sum_{k=1}^{K} D(z^k, T) - D(z, T). \tag{2}$$

$EP(z, z^1, ..., z^K, T)$ in (2) provides a measure of the potential productivity gains (expressed in number of units of the bundle g) obtained from increased specialization. Indeed, assuming that $z^k \neq z/K$ for some k, equation (2) evaluates a change in technical inefficiency (as measured by $D(\cdot)$) comparing two situations: one when netputs z are produced by an integrated firm; and one where netputs z are produced by K "more specialized" firms. D(z, T) in (2) is the distance to the frontier when netputs z are produced in an integrated production process. And $\sum_{k=1}^{K} D(z^k, T)$ is the distance when netputs z are produced in K "more specialized" production processes, z_k being the netputs used in the k-th production process. Given $z = \sum_{k=1}^{K} z^k$, it follows that $EP(z, z^1, ..., z^K, T)$ in (2) has the following interpretation. When $EP(z, z^1, ..., z^K, T) > 0$, the K specialized firms $(z^1, ..., z^K)$ can produce EP additional units of g compared to an integrated firm, implying that specialization improves productivity. It follows that $EP(z, z^1, ..., z^K, T) > 0$ reflects economies of specialization. Alternatively, when $EP(z, z^1, ..., z^K, T) < 0$, the production potential of the K specialized firms $(z^1, ..., z^K)$ is reduced by |EP| units of g compared to an integrated firm, implying that specialization reduces productivity. It follows that $EP(z, z^1, ..., z^K, T) < 0$ reflects diseconomies of specialization.

This is illustrated in Figure 1 which considers the production of two outputs (y_1, y_2) using inputs x, where $z = (-x, y_1, y_2)$. Figure 1 involves a comparison between an integrated firm producing outputs (y_1, y_2) using inputs x and two specialized firms: a firm producing outputs $(y_1, 0)$ using inputs x/2, and a firm producing outputs $(0, y_2)$ using inputs x/2. Figure 1 compares the productivity of the integrated firm producing at point A with the productivity of two specialized firms producing respectively at point C_1 (with netputs $z^1 = (-x/2, y_1, 0)$) and point C_2 (with netputs $z^2 = (-x/2, 0, y_2)$). Note that, as defined, $z = z^1 + z^2$. The evaluation of productivity in Figure 1 relies on the output set $Y(x) \equiv \{(y_1, y_2): (-x, y_1, y_2) \in T\}$. Figure 1 shows that point A is on the frontier of Y(x), with $D(z, T) = 0$. Figure 1 also shows that points C_1 and C_2 are below the frontier of Y(x/2). Given a reference bundle g, the two specialized firms can increase production by the distances $(B_1\ C_1)$ and $(B_2\ C_2)$. From equation (1), these two distances are given by $D(z^1, T)$ and $D(z^2, T)$, respectively. In this case, using (2) and noting that $D(z) = 0$, it follows that $EP(z, z^1, ..., z^K, T) = D(z^1, T) + D(z^2, T) > 0$ measures the potential gain in productivity associated with producing outputs in a specialized manner. Thus, Figure 1 illustrates a situation exhibiting economies of specialization, where specialization increases productivity.

While equation (2) provides a basis to evaluate the gains of specialization, it does not identify where these gains come from. We now explore the sources of these gains. We show next that economies of specialization are closely related to two fundamental concepts: economies of scale and convexity properties of the technology. See the proof in the Appendix.

<u>Proposition 1</u>: At points z and $(z^1, ..., z^K)$ satisfying $z = \sum_{k=1}^{K} z^k$, economies of specialization $EP(z, z^1, ..., z^K)$ in (2) can be decomposed as

$$EP(z, z^1, ..., z^K, T) = ESc(z, T) + ECn(z, z^1, ..., z^K, T), \tag{3}$$

where

$$ESc(z, T) \equiv K\, D(z/K, T) - D(z, T) \tag{4}$$

$$\begin{Bmatrix} \le \\ = \\ \ge \end{Bmatrix} 0 \text{ under } \begin{Bmatrix} \text{increasing returns to scale (IRS)} \\ \text{constant returns to scale (CRS)} \\ \text{decreasing returns to scale (DRS)} \end{Bmatrix},$$

$$ECn(z, z^1, ..., z^K, T) \equiv \sum_{k=1}^{K} D(z^k, T) - K\, D(z/K, T) \tag{5}$$

$$\begin{Bmatrix} \le \\ > \end{Bmatrix} 0 \begin{Bmatrix} \text{if the technology T is convex} \\ \text{only if the technology T is non-convex} \end{Bmatrix}.$$

Equation (3) decomposes economies of specialization $EP(z, z^1, ..., z^K, T)$ into two additive components: the <u>scale component</u> $ESc(z, T)$, and the <u>convexity component</u> $ECn(z, z^1, ..., z^K, T)$. From equation (4), the scale component $ESc(z, T)$ satisfies $ESc(z, T) \begin{Bmatrix} \le \\ = \\ \ge \end{Bmatrix} 0$ under $\begin{Bmatrix} \text{IRS} \\ \text{CRS} \\ \text{DRS} \end{Bmatrix}$. And from equation (5), the convexity component $ECn(z, z^1, ..., z^K, T)$ is always non-positive under a convex technology. It implies that the convexity component $ECn(z, z^1, ..., z^K, T)$ can be positive only under a nonconvex technology.

By definition, <u>diseconomies of specialization</u> exist when $EP(z, z^1, ..., z^K, T) \le 0$. From (3)-(5), this condition always holds under non-decreasing returns to scale (i.e., under either IRS or CRS) and a convex technology. In such situations, there is a disincentive for firms to specialize (i.e., there is an incentive for firms to diversify). Alternatively, economies of specialization exist when $EP(z, z^1, ..., z^K, T) \geq 0$. From (3)-(5), this can arise under DRS and/or under a non-convex technology. This result indicates that DRS and nonconvexity can provide an incentive for firms to specialize.

The decomposition given in equation (3) suggests that both returns to scale and convexity affect the gains from specialization. But are these two factors really different? Indeed, a technology exhibiting IRS is a form of non-convexity for T. If so, is the decomposition presented in (3) really useful? We argue below that it is. Indeed, IRS is form of non-convexity that applies in a very restrictive way: returns to scale consider only proportional changes in all netputs. It means that non-convexity can arise in ways that are unrelated to scale effects. We show below that these forms of non-convexity are important factors that can generate productivity gains from specialization. More fundamentally, as noted in the introduction, we think that the common idea that IRS and non-convexity tend to go together has contributed to hiding the deeper role played by non-convexity. The challenge is to present convincing arguments that returns to scale and non-convexity have different effects on the economies of specialization.

These arguments are illustrated graphically in Figures 2 and 3. These two figures document the independent role of scale versus non-convexity in the evaluation of the gains from specialization. Like Figure 1, Figures 2 and 3 consider the production of two outputs (y_1, y_2) using inputs x, where $z = (-x, y_1, y_2)$. Again, we compare an integrated firm producing outputs (y_1, y_2) using inputs x with two specialized firms: a firm producing outputs $(y_1, 0)$ using inputs x/2, and a firm producing outputs $(0, y_2)$ using inputs x/2. Again, productivity is assessed based on the output set $Y(x) \equiv \{(y_1, y_2): (-x, y_1, y_2) \in T\}$.

Figure 2 evaluates situations where returns to scale can change from IRS to CRS to DRS.

But it holds the convexity of the set Y(x) constant by assuming that the upper bound of the output set Y(x) is linear. The boundary of the output set Y(x) is given by the line A'A" in Figure 2. Assuming that the integrated firm is technically efficient, it produces outputs (y_1, y_2) at point A (located on the line A'A"). The two specialized firms produce outputs (y_1, 0) and (0, y_2), respectively. For these two firms, suppose that the boundary of the output set Y(x/2) is given by the line B_{crs}'B_{crs}" under CRS, by the line B_{irs}'B_{irs}" under IRS, and by the line B_{drs}'B_{drs}" under DRS. Under CRS, the two specialized firms are technically efficient (as both (y_1, 0) and (0, y_2) are located on the frontier B_{crs}'B_{crs}"). In this case, EP(z, z^1, z^2, T) = 0 and there is zero gain from specialization. However, under DRS, the two specialized firms are located at points (y_1, 0) and (0, y_2) that are both below the frontier B_{drs}'B_{drs}". Then EP(z, z^1, z^2, T) > 0, and there are positive potential gains from specialization. Indeed, under DRS, Figure 2 shows that it becomes possible for the two specialized firms to increase aggregate production from point A to point E_{drs}. In contrast, under IRS, the two specialized firms cannot feasibly produce outputs (y_1, 0) or (0, y_2) (as these points are located above the frontier B_{irs}'B_{irs}"). In this case, there are potential negative gains from specialization and EP(z, z^1, z^2, T) < 0 as the aggregate production of the two specialized firms declines from point A to point E_{irs}. Thus, Figure 2 presents a case where economies of specialization can be made positive, zero or negative by just changing the nature of returns to scale. These scale effects are consistent with the results stated in (3) and (4): DRS contributes to economies of specialization, while IRS contributes to diseconomies of specialization.

Figure 3 evaluates situations where the convexity of the output set Y(x) no longer holds: the output set changes from being convex to non-convex. But it assumes CRS, thus holding returns to scale constant. It presents the key argument that economies of specialization can arise for reasons unrelated to scale effects. Again, we start with a technically efficient integrated firm producing outputs (y_1, y_2) at point A, with two specialized firms producing outputs (y_1, 0) and (0, y_2). For the specialized firms, the boundary of the output set Y(x/2) is given in Figure 3 by the line B_c'B_c" when Y(x/2) is convex, by the line B_{lin}'B_{lin}" when the upper bound of Y(x/2) is linear, and by the line B_{nc}'B_{nc}" when Y(x/2) is non-convex. When the upper bound of Y(x/2) is linear, the two specialized firms are technically efficient (as both (y_1, 0) and (0, y_2) are located on the frontier B_{lin}'B_{lin}"). In this case, EP(z, z^1, z^2, T) = 0 and there is zero gain from specialization. Alternatively, under a convex set Y(x), the two specialized firms cannot feasibly produce outputs (y_1, 0) or (0, y_2) (as these points are located above the frontier B_c'B_c"). In this case, there are negative potential gains from specialization and EP(z, z^1, z^2, T) < 0, as the aggregate production of the two specialized firms declines from point A to point E_c. Finally, when Y(x/2) is non-convex, the two specialized firms are located at points (y_1, 0) and (0, y_2) that are both located below the frontier B_{nc}'B_{nc}". Then, EP(z, z^1, z^2, T) > 0 and there are potential productivity gains from specialization. Indeed, under a non-convex Y(x/2), Figure 3 shows that it becomes possible for the two specialized firms to increase aggregate production from point A to point E_{nc}. This illustrates two important points. First, these convexity effects are consistent with the results stated in (3) and (5): convexity contributes to diseconomies of specialization and non-convexity contributes to economies of specialization. Second, Figure 3 presents a case where economies of specialization can be made positive, zero or negative by just changing the convexity properties of the technology, holding returns to scale constant. It means that non-convexity can affect economies of specialization in ways that are unrelated to scale effects. This stresses a need for an in-depth evaluation of the convexity properties of the technology in the evaluation of economies of specialization.

3. Evaluating Economies of Scale and Non-Convexity

The above discussion indicates a need to consider scale effects and the convexity property of the technology together to evaluate economies of specialization. This requires developing empirical methods that can be used for that purpose. The distance function in (1) provides a convenient basis to evaluate the properties of technology.

Definition 2: Let T^h be the convex hull of T, with $T^h \equiv \{\alpha\, z^a + (1-\alpha)\, z^b: z^a \in T, z^b \in T, \alpha \in [0, 1]\} \supset T$. At point z, define the following measure of non-convexity

$$Cn(z) \equiv D(z, T^h) - D(z, T) \geq 0. \qquad (6)$$

The non-negativity of Cn(z) in (6) follows from (1) and $T^h \supset T$. From the definition of convexity, Cn(z) = 0 when the technology T is convex. Alternatively, Cn(z) > 0 implies the presence of non-convexity in T. Thus, Cn(z) in (6) can be interpreted as a local measure (expressed in number of units of g) of the strength of departure from convexity. The measure is local in the sense that it applies at point z.

Definition 3: Let T^c be the cone of T, with $T^c \equiv \{\delta\, z, z \in T, \delta \in R_+\} \supset T$. At point z, define the following measure of economies of scale

$$Sc(z) \equiv D(z, T^c) - D(z, T) \geq 0. \qquad (7)$$

The non-negativity of Sc(z) in (7) follows from (1) and $T^c \supset T$. From the definition of returns to scale, Sc(z) = 0 when the technology exhibits constant returns to scale (CRS). Alternatively, Sc(z) > 0 implies a departure from constant returns to scale. Thus, Sc(z) in (7) can be interpreted as a local measure (expressed in number of units of g) of the strength of departure from CRS. The measure is local in the sense that it applies at point z.

4. Empirical Evaluation of the Technology

The empirical measurements in (6) and (7) require representations of the technology T. Consider a data set involving observations of m netputs chosen by N firms: $z_i = (z_{1i}, \ldots, z_{mi})$, where z_{ji} is the j-th netput used by the i-th firm, $i \in N \equiv \{1, \ldots, N\}$. Following Varian (1984), Färe et al. (1994) and Banker et al. (2004), consider first the following nonparametric representations of technology

$$T_s = \{z: z \leq \Sigma_{i \in N}\ \lambda_i\ z_i;\ \lambda_i \in R_+,\ i \in N,\ \Sigma_{i \in N}\ \lambda_i \in S_s\}, \tag{8}$$

where $s \in \{v, c\}$, with $S_v = 1$ under variable returns to scale (VRS) and $S_c = R^+$ under constant returns to scale (CRS).[6] Under free disposal, T_v in (8) is the smallest convex set containing all data points; and T_c is the smallest convex cone containing all data points. For these reasons, the representations given in (8) have been called "Data Envelopment Analysis" (DEA). Since $S_v \subset S_c$, it follows from (8) that $T_v \subset T_c$. Note that the sets T_v and T_c are both convex.

Next, we want to consider representations of the technology that allow for non-convexity. One possibility is given by the following nonparametric representations

$$T_{FDHs} = \{z: z \leq \Sigma_{i \in N} \lambda_i\ z_i;\ \lambda_i \in \{0, \delta\},\ i \in N;\ \Sigma_{i \in N}\ \lambda_i = \delta;\ \delta \in S_s\}. \tag{9}$$

where FDH stands for "free disposal hull" (Deprins et al., 1984; Kerstens and Eeckaut, 1999), $s \in \{v, c\}$, and the S_s's are as defined above. Again, $S_v = 1$ corresponds to variable returns to scale (VRS) while $S_c = R^+$ corresponds to constant returns to scale (CRS). Under free disposal, T_{FDHv} is the smallest set containing all data points, while T_{FDHc} is the smallest cone containing all data points. Since $S_v \subset S_c$, it follows from (9) that $T_{FDHv} \subset T_{FDHc}$. Note that each of the sets T_v and T_c is in general non-convex. Finally, note that the λ's are restricted to take discrete values in (9) but not in (8). It follows that $T_{FDHs} \subset T_s$, i.e. that T_{FDHs} is a subset of T_s, for $s \in \{v, c\}$.

While the FDH approach allows for non-convexity, it seems to be overly restrictive. Indeed, under VRS, FDH can find non-convexity everywhere marginal products are positive and bounded, which seems too extreme (Chavas and Kim, 2014). Thus, there is a need to develop some alternative and more general characterization of a non-convex technology. For that purpose, define a neighborhood of $z \equiv (z_1, \ldots, z_m) \in R^m$ as $B_r(z, \sigma) = \{z': D_p(z, z') \leq r: z' \in R^m\} \subset R^m$, where $r > 0$ and $D_p(z, z') \equiv \sum_{j=1}^{m} [(|z_j - z_j'|/\sigma_j)^p]^{1/p}$ is a weighted Minkowski distance between z and z', with weights $\sigma = (\sigma_1, \ldots, \sigma_m) \in R^m_{++}$ and based on a p-norm $1 \leq p < \infty$.[7] Following Chavas and Kim (2014), let $I(z, r) = \{i: z_i \in B_r(z, \sigma), i \in N\} \subset N$, where $I(z, r)$ is the set of firms in N that are located in the neighborhood $B_r(z, \sigma)$ of z.[8]

Definition 4: Define a neighborhood-based representation of the technology T as

$$T_{rs}^* = \cup_{i \in N} T_{rs}(z_i), \tag{10}$$

where

$$T_{rs}(z) = \{z: z \leq \Sigma_{i \in I(z,r)} \lambda_i z_i; \lambda_i \in R_+, i \in I(z, r); \Sigma_{i \in I(z,r)} \lambda_i \in S_s\}. \tag{11}$$

with $s \in \{v, c\}$, and the S_s's are as defined above.

The representation of technology given in (10)-(11) is obtained in two steps. In a first step, equation (11) defines $T_{rs}(z)$ as a local representation of the technology T in the neighborhood of point z under free disposal and returns to scale characterized by $s \in \{v, c\}$. Since $S_v \subset S_c$, it follows from (11) that $T_{rv}(z) \subset T_{rc}(z)$. Again, note that, for a given z, the sets $T_{rv}(z)$ and $T_{rc}(z)$ are convex. In a second step, equation (10) defines the set T_{rs}^* as the union of the sets $T_{rs}(z_i)$, $i \in N$. Since the union of convex sets is not necessarily convex, it follows that T_{rs}^*

defined in (10) is not necessarily convex for each $s \in \{v, c\}$. Equation (10) is our proposed neighborhood-based representation of technology. It allows for non-convexity to arise in any part of the feasible set.[9] The properties of T_{rs}^* are presented next. (See the proof in the Appendix).

Proposition 2: For $s \in \{v, c\}$,

$$\lim_{r\to\infty} T_{rs}^* = T_s, \quad (12)$$

$$\lim_{r\to 0} T_{rs}^* = T_{FDHs}, \quad (13)$$

and

$$T_{FDHs} \subset T_{rs}^* \subset T_{r's}^* \subset T_s, \text{ for any } r' > r > 0. \quad (14)$$

Given $s \in \{v, c\}$, equations (12) and (13) show that T_{rs}^* includes two important special cases. From equation (8), the set T_{rs}^* reduces to the set T_s when $r \to \infty$, i.e. when the neighborhood $B_r(z, \sigma)$ of any z becomes "very large". And from equation (13), the set T_{rs}^* reduces to the set T_{FDHs} when $r \to 0$, i.e. when the neighborhood $B_r(z_i, \sigma)$ become "very small" for any $i \in N$.

In addition, equation (14) shows that T_{FDHs} is in general a subset of T_s: $T_{FDHs} \subset T_s$, for $s \in \{v, c\}$. It also establishes that the set T_{rs}^*, our neighborhood-based representation of technology, is bounded between T_{FDHs} and T_s, with T_{FDHs} as lower bound and T_s as upper bound. Noting that the set T_s is convex, and the set T_{FDHs} is in general non-convex, it means that T_{rs}^* provides a generic way of introducing non-convexity in production analysis (Chavas and Kin, 2014). And these representations apply under alternative scale properties: under VRS when $s \in v$ (with S_v = 1), or under CRS when s = c (with $S_c = R^+$). Finally, equation (14) states that the set T_{rs}^* becomes larger when r increases, i.e. when the neighborhoods used to evaluate T_{rs}^* become larger. This provides some flexibility in the empirical analysis of non-convexity issues.

As such, T_{rs}^* has three useful characteristics: 1/ it provides a flexible representation of non-convexity; 2/ it nests as (restrictive) special cases both the DEA model and the FDH model; and 3/ it is easy to implement empirically. To illustrate this last point, given $j \in N$ and $T_{rs}(z_j)$ in (11), note that the evaluation of the distance $D(z, T_{rs}(z_j))$ in (1) involves solving the simple linear programming problem: $D(z, T_{rs}(z_j)) = \text{Max}_{\beta,\lambda} \{\beta: (z + \beta g) \leq \Sigma_{i \in I(z_j,r)} \lambda_i z_i ; \lambda_i \in R_+, i \in I(z_j, r); \Sigma_{i \in I(z_j,r)} \lambda_i \in S_s\}$. And it follows from (10) that $D(z, T_{rs}^*) = \max_i \{D(z, T_{rs}(z_i)): i \in N\}$. As noted above, $D(z, T_{rs}^*)$ is a measure of technical inefficiency (expressed in number of units of the bundle g) for netput z under technology T_{rs}^*. In turn, it provides a basis to evaluate the convexity effect Cn(z) given in (6), with $T^h = T_{\infty s}^*$ and $T = T_{rs}^*$.[10] And it gives a basis to evaluate the scale effects Sc(z) given in (7), with $T^c = T_{rc}^*$ and $T = T_{rv}^*$.

5. Empirical Analysis

The analysis presented above is general: it applies to any firm, irrespective of its institutional form or organization. This section illustrates the usefulness of our approach through an empirical application. The application is to a panel data set of production activities from a sample of Korean rice farms. Focusing on farms is of interest as most farms produce more than one output, allowing us to observed different patterns of specialization across farms. In addition, farms are typically family farms with a simple organizational structure: the head of the household is the manager. And most of the labor is typically provided by family labor, meaning that coordination issues among workers are minimal. To the extent that labor and management are often performed for the same person, we can expect the gains from specialization to be closely associated with the managerial skills of the farm manager, i.e. his ability to manage multiple farm production activities. Our empirical analysis will evaluate the nature of scale effects Sc(·) given in (6) and non-convexity effects Cn(·) given in (7). In turn, we will examine the factors contributing to non-convexity.

5.1. Data

The analysis uses farm household level data from Korea, data collected in a Farm Household Economy Survey between 2003 and 2007 by the National Statistical Office (Kim et al., 2012). This annual survey provides data on the farm household economy and agricultural management. The data come from a sample of 3,140 farm households surveyed annually from 314 enumeration districts. These districts are sampled first using a proportional sampling scheme based on the number of farm households from Agricultural Census at 2000. Although this survey includes 8 different farm household types which are determined by the largest proportion of the farm household revenue including paddy rice farming and vegetables farming, our empirical analysis focuses on a sample of farms classified as “paddy rice farms” located in the Jeon-Nam province in the southern part of Korea. While most farms produce more than one output, the farms in our sample have a relatively high share of farm revenue coming from rice. The reason why we focus only on rice farms in Jeon-Nam province is that this area has an extensive irrigation network supporting rice production and is known as a rice-producing province. Moreover, being in the same region, it is relatively safe to assume that all farms face similar agro-climatic conditions. The sample includes 86, 120, 101, 101, 122 number of rice farms for the year of 2003, 2004, 2005, 2006, and 2007, respectively. This unbalanced panel dataset contains data on nine netputs: four outputs and five inputs. The outputs are: rice, vegetable, livestock and other outputs. The inputs are family labor, paddy land owned, non-paddy land owned, land rented, and other inputs. Family labor input is measured in hours, and land inputs are measured in hectares (ha). Other netputs are measured in value, assuming that all farmers face the same prices. Summary statistics are presented in Table 1. The average revenue from rice production is 14,990.5 (measured in 1,000 won[11]), accounting for 64.2% of total farm revenue. The second largest source of revenue is vegetable production: 3,177.0 (measured in 1,000 won), accounting for 15.1% of total farm revenue. The average size of a farm is 2.58 ha (including both land owned and land rented). The sample reflects the type of farms commonly found in Asia where farms are typically small and with some degree of specialization in rice production.

5.2. Results

The analysis relies on nonparametric representations of the technology T_{rs}^* given in (10). The distance function D(z, T) in (1) is evaluated based on the bundle $g = (g_1, \ldots, g_m)$ such that $g_i = 0$ for the i-th input and g_j is the sample mean for the j-th output. Thus, our reference bundle $g = (g_1, \ldots, g_m)$ is the typical bundle associated with the outputs of an average farm. This choice leads to a simple interpretation of our directional distance estimates. For example, for a given T, finding that $D(z_i, T) = 0.2$ means that the i-th farm is technically inefficient: it could move to the production frontier and increase its outputs by 20 percent of the average outputs in the sample. This interpretation remains valid under alternative characterizations of the technology T.

Our neighborhood-based assessment of technology T_{rs}^* requires the definition of a neighborhood. Letting $M_j = \mathrm{Max}_{i \in N} \{z_{ji}\} - \mathrm{Min}_{i \in N} \{z_{ji}\}$ be the sample range of the j-th netput, we considered dividing the sample range into four equally spaced intervals and defined neighborhoods as $B_r(z, \cdot) = \{z': -M_j/4 \le z_j - z_j' \le M_j/4;\ j = 1, \ldots, m;\ z' \in R^m\}$.[12] Based on these neighborhoods, our empirical analysis generates farm-specific estimates of technical inefficiency measured by $D(z_i, T_{rs}^*)$, $i \in N$. And it permits an evaluation of farm-specific convexity effects

$Cn(z_i)$ given in (6) and of scale effects $Sc(z_i)$ given in (7), $i \in N$.

The empirical analysis uses annual data on production activities from a sample of 530 Korean farms over the period 2003-2007. Our results will be evaluated for three farm types: small farms, medium farms and large farms. Small farms are defined as farms located in the lower 30 percentile distribution of total land; large farms are those farms located in the top 30 percentile distribution of total land; and medium farms are in between.

First, we evaluated the convexity effect Cn given in (6), with $T^h = T_{\infty s}{}^*$ and $T = T_{rs}{}^*$. The results are summarized in Table 2. Table 2 presents average values of Cn for each year (2003, 2004, 2005, 2006 and 2007), for each farm type (small, medium and large farms), and under both CRS (s = c) and VRS (s = v). The results show that Cn varies between 0.007 and 0.177. This documents that the non-convexity effects can be large. For example, Cn = 0.177 means that non-convexity effects accounts for a 17.7 percent change in the mean value of all outputs. The estimates of Cn are fairly similar for CRS versus VRS, indicating that the presence of non-convexity is not related to scale effects. In general, Table 2 shows that Cn tends to be moderate for small farms (always less than 0.03) but that they increase with farm size. Indeed, with the exception of (2005, CRS), the largest Cn estimates are consistently found among large farms. This provides evidence that non-convexity effects become stronger on larger farms. It means that specialized operators tend to be more productive on larger farms. To the extent that non-convexity comes from the saving in fixed cost related to labor and managerial resources, this would imply that the productivity of specialized management improve more on large farmers. Finally, Table 2 shows that some changes in the Cn estimates over time, although not clear patterns seem to emerge. This is consistent with a slow technology change in rice production in Korea, reflected by a complete irrigation infrastructure and high-yielding rice varieties available by farmers.

Second, we evaluated the scale measure Sc given in (7), with $T^c = T_{rc}^*$ and $T = T_{rv}^*$. The results are summarized in Table 3. Table 3 presents average values of Sc for each year (2003, 2004, 2005, 2006 and 2007), for each farm type (small, medium and large farms), and under both convexity (when $r \to \infty$) and non-convexity. The results show that Sc varies between 0.009 and 0.124. With the exception of (convexity, large farms, 2007), the Sc's are all below 0.10. This documents that scale effects are present, but that the magnitude of scale inefficiency is moderate. For example, Sc = 0.009 (convexity, medium farm, 2004) means that scale inefficiency amounts to a 0.9 percent change in the mean value of all outputs. As might be expected, Table 3 shows that most of the scale inefficiency is due to IRS for small farms, but DRS for large farms. As discussed in section 2, we expect DRS (IRS) to contribute positively (negatively) to economies of specialization. As a result, scale effects would provide incentives for large farms to specialize. Yet, the relative small magnitudes of Sc indicate that scale effects are moderate, which is consistent with long-lasting small-scale rice farming in Korea. And they tend to be smaller than the non-convexity effects Cn reported in Table 2. It generates one of our key findings: convexity effects tend to dominate scale effects in Korean agriculture. In other words, while scale effects can affect the gains from specialization, our results point to a dominant role played by non-convexity. From table 3, these results seem to hold both under a convex technology and a non-convex technology, indicating that scale effects appear to be unrelated to non-convexity. Finally, Table 3 shows some changes in the Sc estimates over time, although not clear patterns seem to emerge. This may reflect slow technology change in rice production in Korea.

Next, we examined the factors associated with non-convexity. Using our farm-specific estimates of Cn, we regressed them on selected explanatory variables. Since the Cn's have zero as a lower bound, we use censored/Tobit regression. If non-convexity is associated with the saving in fixed cost related to labor and managerial resources, then the rise of non-convexity would likely be linked with human capital. On that basis, we include age, schooling, and their interactions as explanatory variables in the Tobit model. And to examine whether non-convexity varies with size, we also include farm size as an explanatory variable. Summary statistics for these variables are presented in Table 4.

We estimated the Tobit model using the Cn estimates obtained under CRS as well as VRS. The Tobit estimates are reported in Table 5. The estimates show that age, schooling and their interaction have each statistically significant effects on non-convexity. This is consistent with our interpretation of non-convexity being associated with saving in fixed cost related to labor and managerial resources. Our results indicate that managerial ability likely changes with both experience and education. Interestingly, due to the interaction effects, the marginal impacts of age or schooling on non-convexity can be either positive or negative. The marginal impact of age is found to be negative but only for "low schooling." Similarly, the impact of schooling is found to be negative but only for young individuals. Evaluated at sample mean of schooling, we found a negative relationship between age and non-convexity. This indicates that younger individuals have more incentive to specialize in rice production holding other variables constant. Evaluated at sample mean of age, we found a positive relationship between schooling and non-convexity. This implies that education contributes to specialization in rice production. To the extent that non-convexity contributes to gains from specialization, our results indicate that young and better educated individuals would have more incentive to specialize. Given the fact that larger rice farms are generally operated by relatively younger individuals seeking specialization benefits associated with rice farming by increasing the size of paddy land, this result seems plausible in a Korean rice production context. Alternatively, older and less educated individuals would have less incentive to specialize. This suggests that the pattern of specialization would vary with education and with the life cycle of individuals. The Tobit estimates reported in Table 5 also show that farms that are more specialized in rice exhibit more non-convexity. This likely reflects the presence of significant fixed cost associated with rice production. Finally, Table 5 shows that farm size has a strong and positive relationship with non-convexity. This is consistent with the results reported in Table 2: non-convexity effects tend to be more important on large farms. Again, we interpret this result as evidence that the productivity of specialized management improve more on larger farmers.

Finally, simulations of economies of specialization (EP given in (2)) and its scale component (ESc given in (4)) and convexity component (ECn given in (5)) are presented in Table 6 for selected farm types. Two farm types are evaluated: a moderate-size farm with 1.77 ha of land, and a large farm size with 5.23 ha of land. The simulation involves two specialization schemes (K = 2), with z^1 being specialized in rice and z^2 being specialized in other (non-rice) activities.[13] Table 6 reports economies of specialization for the large farm (with EP = 0.608), but diseconomies of specialization for the moderate-size farm (with EP = -0.262). Importantly, these specialization effects are large. For example, the reference bundle g representing average farm outputs in the sample, EP = 0.608 measures productivity effects amounting to a 60.8 percent increase in average revenue. This illustrates that the incentives to specialize are strong on large farms but not on smaller farms. Table 6 also shows that the scale component ESc is negative for both farm types (ESc = -0.091 for moderate size, and ESc = -0.137 for large farm). From equation (4), this corresponds to situations of increasing returns to scale (IRS), which tends to reduce the benefit of specialization. Finally, Table 6 shows that the convexity component ECn is negative for the moderate-size farm (ECn = -0.171), but positive for the large farm (ECn = 0.745). As stated in equation (5) above, ECn is necessarily negative under a convex technology; and it can turn positive only in the presence of non-convexity. Thus, Table 6 shows two important results: 1/ the productivity benefits of specialization come in large part from non-convexity; and 2/ non-convexity is prevalent in large farms and not in smaller farms. This documents the role of non-convexity and its effects on the incentive for farms of different types to specialize. In a way consistent with our earlier results, it reveals significant productivity benefits from specialized management on larger farms, providing a strong incentive for large farms to specialize.

6. Concluding Remarks

This paper has presented an analysis of the microeconomics of firm specialization. We have proposed a measure of economies of specialization, reflecting the productivity effects of greater firm specialization. We have identified the distinct role played by returns to scale versus non-convexity. Our conceptual analysis showed that diseconomies of firm specialization occur under increasing returns to scale and a convex technology. Alternatively, economies of firm specialization arise under decreasing returns to scale and a non-convex technology. This indicates that a need for a combined empirical assessment of the nature of returns and scale and non-convexity. In this context, we developed measures of economies of scale and non-convexity and proposed methods to evaluate them empirically. The usefulness of the approach was illustrated in an empirical application to a data set of Korean farms. The analysis documented the presence of non-convexity as well as scale effects. We showed that non-convexity varies across farm types. Non-convexity was found to be more common on larger farms, indicating that specialized operators have a greater ability to improve productivity on larger farms. We also found that non-convexity effects are more important than scale effects on larger farms. This has two implications. First, it means that scale effects are not the major factor affecting farm specialization. Second, the changes in non-convexity effects across farm size can help explain why larger farms tend to be more specialized. Our empirical analysis also evaluates the linkages between management and non-convexity. Most farms being family farms, we find that non-convexity varies with the education and experience of the farm manager. This provides new insights into the role of management and its implications for firm productivity and the economics of specialization.

Our analysis could be extended in several directions. First, the gains of specialization need to be analyzed at the aggregate level. This means examining how the micro effects analyzed in this paper translate into macro effects. This should provide new insights into the aggregate gains from trade. In particular, our analysis indicates that the gains from trade likely vary across industries depending on the convex (or non-convex) nature of the technology. Second, our application has focused on agriculture. There is a need to expand the empirical analysis to other industries. Third, while we documented that the linkages between specialization and a non-convex technology, there is a need for further empirical studies of the linkages between management and specialization gains. Exploring these issues appears to be good topics for future research.

References

Agrell, P.J., P. Bogetoft and J. Tind. "Efficiency Evaluation with Convex Pairs. *Advanced Modeling and Optimization*, 7(2005): 211-237.

Arkolakis, C. A. Costinot and A. Rodriguez-Clare. "New Trade Models, Same Old Gains?" *American Economic Review* 102(2012): 94-130.

Balistreri, E., J., R.H. Hillberry and T.F. Rutherford. "Structural Estimation and Solution of International Trade Models with Heterogeneous Firms" *Journal of International Economics* 83(2011): 95-108.

Banker, R.D., W.W. Cooper, L.M. Seiford, R.M. Thrall and J. Zhu. "Returns to Scale in Different DEA Models" *European Journal of Operation research* 154(2004): 345-362.

Baumol, W.J., J.C. Panzar and R. Willig. *Contestable Markets and the Theory of Industry Structure*, Harcourt Brace Jovanovich, Inc., New York, 1982.

Becker, G.S. and K.M. Murphy, "The Division of Labor, Coordination Costs and Knowledge" *Quarterly Journal of Economics* 107(1992): 1137-1160.

Bernard, A.B., J. Eaton, J.B. Jensen and S. Kortum. "Plants and Productivity in International Trade" *American Economic Review* 93(2003): 1268-1290.

Bogetoft, P. "DEA on Relaxed Convexity Assumptions" *Management Science* 42(1996): 457-465.

Caliendo, L. and E. Rossi-Hansberg. "The Impact of Trade on Organization and Productivity" *Quarterly Journal of Economics* 127(2012): 1393-1467.

Chambers, R.G., Y. Chung and R. Färe. "Benefit and Distance Functions" *Journal of Economic Theory* 70 (1996): 407-419.

Chavas, J.P. "Structural Change in Agricultural Production: Economics, Technology and Policy" in *Handbook in Agricultural Economics, Vol. 1*, B. Gardner and G. Rausser, editors, Elsevier Science (2001): 263-285.

Chavas, J.P. and K. Kim. "Measurement and Sources of Economies of Scope" *Journal of Institutional and Theoretical Economics* 163(2007): 411-427.

Chavas, J.P. and K. Kim. “Nonparametric Analysis of Technology and Productivity under Non-Convexity: A Neighborhood-Based Approach” *Journal of Productivity Analysis* (2014): forthcoming.

Coviello, D., A. Ichino and N. Persico. “Time Allocation and Task Juggling” *American Economic Review* 104(2014): 609-623.

Deprins, D., L. Simar and H. Tulkens. “Measuring Labor Efficiency in Post Offices” In: M. Marchand, P. Pestiau and H. Tulkens (eds). *The Performance of Public Enterprises: Concepts and Measurements*. North-Holland: Amsterdam (1984): 243-267.

Färe, R., S. Grosskopf and C.A.K. Lovell. *Production Frontiers* Cambridge University Press, Cambridge, 1994.

Färe, R. and S. Grosskopf. “Theory and Application of Directional Distance Functions” *Journal of Productivity Analysis* 13(2000): 93-103.

Farrell, M. “The Measurement of Productive Efficiency” *Journal of the Royal Statistical Society*. Series A, General, Part 3, 120(1957): 253-281.

Feenstra, R.C. “Estimating the Effects of Trade Policy” in *Handbook of International Economics,* Vol. 3, Edited by G. Grossman and K. Rogoff, Elsevier Science (1995): 1553-1595.

Kim, K., J.P. Chavas, B. Barham and J. Foltz. "Specialization, Diversification, and Productivity: A Panel Data Analysis of Rice Farms in Korea" *Agricultural Economics* 43(2012): 1-14.

Kerstens, K. and P.V. Eeckaut. “Estimating Returns to Scale Using Non-Parametric Deterministic Technologies: A New Method Based on Goodness-of-fit” *European Journal of Operation Research* 113(1999): 206-214.

Krugman, P. “Scale Economies, Product Differentiation and the Patterns of Trade” *American Economic Review* 70(1980): 950-959.

Luenberger, D. G. *Microeconomic Theory* McGraw-Hill, Inc., New York, 1995.

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

Melitz, M.J. and D. Trefler. "Gains from Trade when Firms Matter" *Journal of Economic*

Perspectives 26(2012): 91-118.

Melitz, M.J. and S.J. Redding. "Missing Gains from Trade**?"** *American Economic Review* 104(2014): 317-321.

Petersen, N.C. "Data Envelopment Analysis on a Relaxed Set of Assumptions" *Management Science* 36(1990): 214-305.

Ricardo, D. *On the Principles of Political Economy and Taxation* (John Murray, London, 1817), in *The Works and Correspondence of David Ricardo*, Vol. 1, ed. Piero Sraffa, Cambridge University Press, Cambridge, 1951.

Samuelson, P. "The Gains from International Trade Once Again" *Economic Journal* 72(1962): 820-829.

Shephard, R.W. *Cost and Production Functions*. Princeton University Press, Princeton, 1953.

Smith, A. *The Wealth of Nations* (London, 1776), Hackett Publishing Company, Indianapolis, 1993.

Stigler, G.J. "The Division of Labor is Limited by the Extent of the Market" *Journal of Political Economy* 59(1951): 185-193.

Varian, H.R. "The Nonparametric Approach to Production Analysis" *Econometrica* 52(1984): 579-597.

Appendix

Proof of Proposition 1: Equation (2) can be alternatively written as $EP(z, z^1, ..., z^K, T) = \sum_{k=1}^{K} D(z^k, T) - K\, D(z/K, T) + K\, D(z/K, T) - D(z, T)$. Given (4) and (5), this gives the decomposition in (3).

For K > 1, note that $T \begin{Bmatrix} \supset \\ = \\ \subset \end{Bmatrix} K\, T$ under $\begin{Bmatrix} \text{IRS} \\ \text{CRS} \\ \text{DRS} \end{Bmatrix}$. It follows from (1) that $D(z, T) = \sup_\beta \{\beta: (z + \beta\, g) \in T\} \begin{Bmatrix} \geq \\ = \\ \leq \end{Bmatrix} \sup_\beta \{\beta: (z + \beta\, g) \in K\, T\}$ under $\begin{Bmatrix} \text{IRS} \\ \text{CRS} \\ \text{DRS} \end{Bmatrix}$. Letting $b = \beta/K$, we have $\sup_\beta \{\beta: (z + \beta\, g) \in K\, T\} = K \sup_b \{b: (z/K + b\, g) \in T\} = K\, D(z/K, T)$. Combining these results gives the inequalities in (4).

From (1), we have $D(z^k, T) = \sup_\beta \{\beta_k : (z^k + \beta_k\, g) \in T\}$ and $\sum_{k=1}^{K} (1/K)\, D(z^k, T) = \sup_\beta \{\sum_{k=1}^{K} \beta_k/K: (z^k + \beta_k\, g) \in T, k = 1, \ldots, K\}$. Assume that the set T is convex. Then, $(z^k + \beta_k\, g) \in T$ for all k implies that $\sum_{k=1}^{K} [z^k/K + (\beta_k/K)\, g] \in T$. Letting $\alpha = \sum_{k=1}^{K} \beta_k/K$, it follows that $\sup_\beta \{\sum_{k=1}^{K} \beta_k/K: (z^k + \beta_k\, g) \in T, k = 1, \ldots, K\} \leq \sup_\alpha \{\alpha: (\sum_{k=1}^{K} z^k/K + \alpha\, g) \in T\} = D(\sum_{k=1}^{K} z^k/K, T)$. When $z = \sum_{k=1}^{K} z^k$, this yields $\sum_{k=1}^{K} (1/K)\, D(z^k, T) \leq D(z/K, T)$, which gives the first inequality in (5).

Proof of Proposition 2:

Note that $\lim_{r\to\infty} I(z, r) = N$ for any finite $z \in R^m$. It follows from (8), (10) and (11) that $T_s = \lim_{r\to\infty} T_{rs}(z_i) = \lim_{r\to\infty} T_{rs}^*$ for any $i \in N$ and $s \in \{v, c\}$. This gives (12).

We have $\lim_{r\to 0} B_r(z_i, \sigma) = \{z_i\}$ and $\lim_{r\to 0} I(z_i, r) = \{i\}$ for any $i \in N$. Using equation

(11), it follows that $\lim_{r\to 0} T_{rs}(z_i) = \{z: z \leq \gamma\, z_i,\ \gamma \in S_s\}$. Equation (10) can be alternatively written as $T_{rs}^{*} = \{\Sigma_{i\in N}\, \alpha_i\, T_{rs}(z_i): \alpha_i \in \{0, 1\},\ i \in N;\ \Sigma_{i\in N}\ \alpha_i = 1\}$. Letting $\eta_i = \alpha_i\, \gamma$, this implies that $\lim_{r\to 0} T_{rs}^{*} = \{z: z \leq \Sigma_{i\in N}\, \eta_i\, z_i;\ \eta_i \in \{0, \gamma\},\ i \in N;\ \Sigma_{i\in N}\ \eta_i = \gamma,\ \gamma \in S_s\}$. Using equation (9), this gives (13).

We have $\lim_{r\to 0} B_r(z_i, \sigma) \subset B_r(z_i, \sigma) \subset B_{r'}(z_i, \sigma) \subset \lim_{r\to\infty} B_r(z_i, \sigma)$ for any $r' > r > 0$. Thus, for any $r' > r > 0$, $\lim_{r\to 0} I(z_i, r) \subset I(z_i, r) \subset I(z_i, r') \subset \lim_{r\to\infty} I(z_i, r) = N$. Then, equation (11) implies that $\lim_{r\to 0} T_{rs}(z_i) \subset T_{rs}(z_i) \subset T_{r's}(z_i) \subset \lim_{r\to\infty} T_{rs}(z_i)$ for any $r' > r > 0$ and any $i \in N$. Using equations (10), (12) and (13), this proves (14).

Table 1. Descriptive statistics

Variables	Number of observations	Sample mean	Standard deviation	Minimum	Maximum
rice revenue (in 1,000 won)	530	14990.6	19413.5	603.7	161260.9
vegetable revenue (in 1,000 won)	530	3176.8	4724.3	0	39649.2
livestock revenue (in 1,000 won)	530	1659.2	3383.1	0	24517.0
other revenue (in 1,000 won)	530	3114.9	5725.3	0	73816.2
production costs (in 1,000 won)	530	10185.9	10763.6	617.5	72654.9
family labor (hours)	530	891.9	565.9	71.5	3634.6
paddy land owned (in ha)	530	1.09	1.40	0	13.52
land owned except paddy land owned (in ha)	530	0.48	0.66	0	6.13
land rented (in 1,000 won)	530	1.01	1.51	0	16.37

Note: 1,000 won (the Korean currency) is approximately equivalent to 0.89 US dollar.

Table 2. Average Non-Convexity Effects Cn(·) under CRS and VRS, by Farm Size over Time

	2003	2004	2005	2006	2007
Under CRS (with $S_c = R^+$)					
Small farm[a]	0.021	0.012	0.017	0.016	0.007
Medium farm	0.026	0.058	0.069	0.051	0.036
Large farm	0.103	0.143	0.067	0.177	0.113
Under VRS (with $S_v = 1$)					
Small farm	0.015	0.007	0.010	0.008	0.006
Medium farm	0.030	0.062	0.068	0.053	0.038
Large farm	0.051	0.143	0.086	0.144	0.080

Note: a/ Farm size is identified by the size of total land. Small farms are defined as farms being in the 0 to 30 percentile of the sample distribution of farm size, medium farms are between the 30 percentile and 70 percentile, and large farms are in the 70 to 100 percentile.

Table 3. Average Scale Effects Sc(·) under Alternative Representations of the Technology, by Farm Size over Time

	2003	2004	2005	2006	2007
Under Convexity: T_v versus T_c					
Small farms[a]					
Average Sc(·)	0.021	0.017	0.021	0.017	0.017
Sc(·) due to IRS	0.018	0.017	0.020	0.016	0.015
Sc(·) due to DRS	0.003	0.0000	0.001	0.001	0.002
Medium farms					
Average Sc(·)	0.041	0.009	0.033	0.020	0.023
Sc(·) due to IRS	0.002	0.005	0.013	0.015	0.007
Sc(·) due to DRS	0.039	0.004	0.020	0.005	0.016
Large farms					
Average Sc(·)	0.084	0.047	0.031	0.089	0.124
Sc(·) due to IRS	0.001	0.002	0.007	0.001	0.001
Sc(·) due to DRS	0.083	0.045	0.024	0.088	0.123
Under Non-Convexity: T_{rv}^* versus T_{rc}^*					
Average Sc(·)					
For small farms	0.015	0.011	0.014	0.009	0.017
For medium farms	0.046	0.013	0.032	0.023	0.025
For large farms	0.032	0.046	0.050	0.056	0.090

Note: a/ Farm size is identified by the size of total land. Small farms are defined as farms being in the 0 to 30 percentile of the sample distribution of farm size, medium farms are between the 30 percentile and 70 percentile, and large farms are in the 70 to 100 percentile.

Table 4. Descriptive Statistics for the Analysis of Non-Convexity

Variable	**Obs.**	**Sample mean**	**Std. deviation**	**Min.**	**Max.**
Non-convexity	530	0.060	0.138	0.000	1.201
Age	530	63.29	9.32	39.00	85.00
Years of schooling	530	6.78	3.77	0.00	16.00
Farm size	530	2.57	2.69	0.18	22.62
Time trend	530	3.10	1.41	1.00	5.00
Rice revenue ratio	530	0.64	0.16	0.28	1.00

Table 5. Tobit Estimation of Factors Affecting Non-convexity (Dependent variable = Cn)

Variables	(a) Under CRS		(b) Under VRS	
	Coefficients	Standard errors	Coefficients	Standard errors
Intercept	0.188	0.186	0.086	0.212
Age	-0.006**	0.003	-0.006**	0.023
Schooling	-0.049**	0.020	-0.048**	0.023
Age*Schooling	0.001**	0.000	0.001**	0.000
Farm size	0.014***	0.005	0.007	0.005
Time trend	0.004	0.008	0.004	0.010
Rice revenue ratio	0.202***	0.075	0.263***	0.087
Sigma	0.230	0.011	0.252	0.014

Note: Stars denote the significance level: *** for the 1 percent significance level; ** for the 5 percent significance level; and * for the 10 percent significance level.

Table 6. Simulations of EP, ESc and ECn for Selected Farm Types[a/]

	Economies of specialization EP	Scale component ESc	Convexity component ECn
	$EP = \sum_k D(z^k, T) - D(z, T)$	$ESc = K\, D(z/K, T) - D(z, T)$	$ECn = \sum_k D(z^k, T) - K\, D(z/K, T)$
Moderate-size farm	-0.262	-0.091	-0.171
Large farm	0.608	-0.137	0.745

Note: The farm size is 1.77 ha for a moderate-size farm and 5.23 ha for a large farm. The simulated specialization schemes are: K = 2, with z^1 being specialized in rice and z^2 being specialized in other (non-rice) activities.

Figure 1: Evaluating the benefit of specialization: the case of two outputs $(y_1, y_2) \in Y(x)$

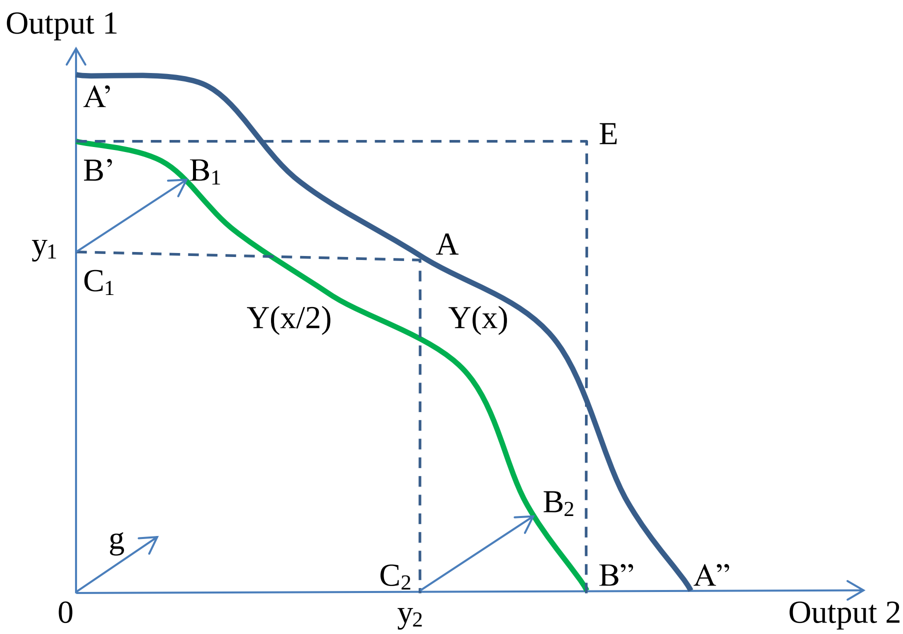


Figure 2: Evaluating the benefit of specialization: the role of returns to scale, $(y_1, y_2) \in Y(x)$

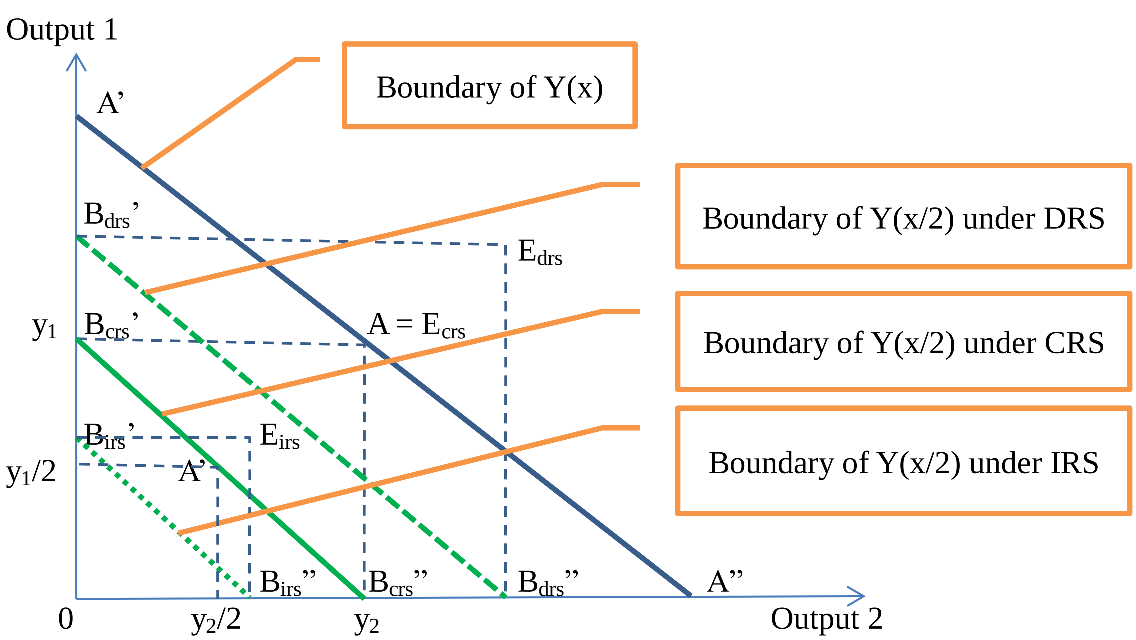
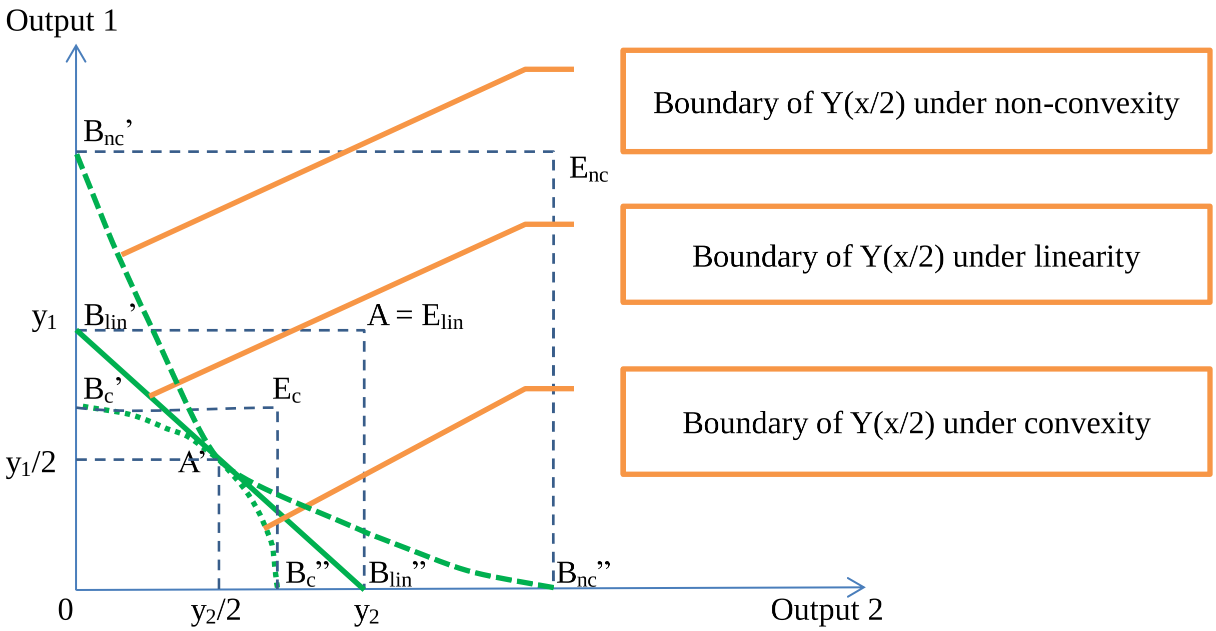


Figure 3: Evaluating the benefit of specialization: the role of convexity, $(y_1, y_2) \in Y(x)$



Footnotes

[1] While our analysis focuses at firm level, there is an extensive literature on the aggregate benefits of specialization. The two approaches (micro versus macro) are obviously related. By arguing that the division of labor is "limited by the extent of the market", Smith (1676, p. 13-17) first stressed that the benefits of specialization can be obtained only in the presence of exchange (see also Stigler (1951)). In this context, economists have shown that the aggregate gains from trade are positive (e.g., Ricardo, 1817; Samuelson, 1962). Yet, some controversy remains about the magnitude of these gains. The empirical measurements of aggregate gains from trade have typically been relatively small. For example, Arkolakis et al. (2012, p. 95) have estimated that the welfare gains from trade for the US have ranged from 0.7% to 1.4% of income. This has stimulated the search for new models that could generate larger gains from specialization and trade. In particular, the roles of economies of scale, product differentiation, imperfect competition and firm heterogeneity have been examined (e.g., Krugman, 1980; Melitz, 2003; Bernard et al., 2003; Balistreri et al., 2011; Melitz and Trefler, 2012; Caliendo and Rossi-Hansberg, 2012; Melitz and Redding, 2013). Yet, Arkolakis et al. (2012) showed that these new inquiries have not had much of an impact on the aggregate gains from specialization and trade. By focusing at the micro level, our paper does not examine the macro side of gains from specialization. Yet, we see our analysis as an important building block toward a better understanding of the economics of specialization at any level.

[2] This paper is a follow-up to Kim et al. (2012). While Kim et al. (2012) relied on parametric method, this paper uses a more flexible nonparametric method to investigate the economics of specialization. On stressed below, our analysis provides new insights on the role of management in specialization incentives.

[3] The directional distance function D(z, T) in (1) is the negative of Luenberger's shortage function (see Luenberger, 1995).

[4] Note that D(z, T) includes as special cases many measures of technical inefficiency that have appeared in the literature. Relationships with Shephard's distance functions (Shephard, 1953) or Farrell's measure of technical efficiency (Farrell, 1957) are discussed in Chambers et al. (1996) and Färe and Grosskopf (2000).

[5] In the special case where the set T is convex, then the distance function D(z, T) is concave in z (Luenberger, 1995).

[6] Note that (8) could be modified to distinguish between IRS, CRS or DRS. Indeed, replacing $[\Sigma_{i \in N}\ \lambda_i \in S_s]$ in (8) by $\Sigma_{i \in N}\ \lambda_i \leq 1\ (\geq 1)$ would correspond to imposing non-increasing returns to scale (non-decreasing returns to scale).

[7] For example, when p = 2, this corresponds to the Euclidean distance: $D_2(z, z') \equiv \sum_{j=1}^{m} [(|z_j - z_j'|/\sigma_j)^2]^{1/2}$. And when $p \to \infty$, this corresponds to the Chebyshev distance: $\lim_{p\to\infty} D_p(z, z') = \text{Max}_j \{|z_j - z_j'|/\sigma_j: j = 1, \ldots, m\}$.

[8] The choice of the neighborhood $B_r(z, \sigma)$ is further discussed in section 4 below.

[9] Nonparametric analyses of non-convex technology have been previously analyzed by Petersen (1990), Bogetoft (1996), Agrell et al. (2005) and Podinovski (2005). The relationships between our approach and previous analyses are discussed in Chavas and Kim (2014).

[10] Noting that $\lim_{r\to\infty} T_{rs}^* = T_s$ from (12), it follows that $D(z, T_{\infty s}^*)$ can be obtained more directly as $D(z, T_s) = \text{Max}_{\beta,\lambda} \{\beta: (z + \beta\ g) \leq \Sigma_{i \in N}\ \lambda_i\ z_i\ ;\ \lambda_i \in R_+,\ i \in I;\ \Sigma_{i \in N}\ \lambda_i \in S_s\}$.

[11] Note that 1,000 won (the Korean currency) = 0.89 US dollars.

[12] We also conducted the analysis based of alternative choices of neighborhoods. As expected (from Proposition 2), we found that choosing smaller (larger) neighborhoods contributed to uncovering more (less) evidence of non-convexity. The results are available upon requests.

[13] In the simulation, the specialized netputs z^1 and z^2 are defined as follows. Compared to the original farm (z), the farm specialized in rice (z^1) produces 70 percent of the rice output, 30 percent of the non-rice outputs, and 50 percent of inputs. And compared to the original farm, the farm specialized in non-rice (z^2) produces 30 percent of the rice output, 70 percent of the non-rice outputs, and 50 percent of inputs. In a way consistent with equation (2), this guarantees that $z = z^1 + z^2$. We chose this pattern of partial output specialization as no farm in our sample was observed to be completely specialized (i.e., producing only rice or only non-rice outputs).