J Prod Anal
DOI 10.1007/s11123-012-0327-6

Returns to scope: a metric for production synergies demonstrated for hospital production

Brandon Pope · Andrew Johnson

Abstract Knowledge of the production function's scope properties can provide insights for firms choosing their operating strategy, policy-makers considering industry structure, and analysts determining appropriate tools. We introduce a new property, returns to scope, which is distinct from scale properties and does not rely on price information. Based on desirable characteristics of an estimator of returns to scope, we propose two methods for assessment. We present examples using simulated data and hospital production data from the 2008 National Inpatient Sample of the Agency for Healthcare Research and Quality's Healthcare Cost and Utilization Project. We find that hospitals experience negative returns to scope (productivity losses) from the joint production of minor and major diagnostic procedures. Based upon our results we conclude that the new returns to scope property allows sharper insights than classic economies of scope approaches.

Keywords Returns to scope · Diversification · Hospital production

JEL Classification D24 · I12

1 Introduction

Scope properties of production functions describe how multiple outputs are traded off against each other, and scale properties describe the trade-offs between inputs and outputs. Despite the relative shortage of theoretical attention in the productivity literature, scope properties have been applied in industries including agriculture (Paul and Nehring 2005), transportation (Rawley and Simcoe 2010; Growitsch and Wetzel 2009), healthcare (Preyra and Pink 2006), education (Sav 2004), banking (Ferrier et al. 1993), R&D (Henderson and Cockburn 1996; Arora et al. 2009), semiconductor manufacturing (Macher 2006), and telecom (Evans and Heckman 1984). These applications typically focus on the classic notions of economies and diseconomies of scope (Panzar and Willig 1981) and their derivatives. The applied interest in scope properties is largely due to its utility in strategic decision-making related to product mix, mergers, outsourcing, and diversification. Extending the literature, this paper introduces returns to scope, a new scope property with several appealing features such as disentanglement from scale properties and non-reliance on price information. We give two methods which assess returns to scope from production data, and illustrative examples using simulated as well as hospital production data from the 2008 National Inpatient Sample of the Agency for Healthcare Research and Quality (AHRQ) Healthcare Cost and Utilization Project (HCUP). Our results can be interpreted as giving insight as to the potential productivity gains (or losses) from joint production of outputs. For example, we find that hospitals of all sizes experience negative returns to scope (productivity losses) from the joint production of minor and major diagnostic procedures.

The remainder of this paper is organized as follows. Section 2 reviews the literature on scope properties of production functions and their assessment. In Sect. 3 we introduce the new property, returns to scope, and examine its relationship to economies of scope. Section 4 discusses the desirable properties of an assessment of returns to scope, and presents two such methods. Section 5 describes

B. Pope (✉)
School of Industrial Engineering, Purdue University, 315 N. Grant Street, West Lafayette, IN 47907-2023, USA
e-mail: popeb@purdue.edu

A. Johnson
Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA

Published online: 23 November 2012

an example using simulated data, and then applies the new property and methods to hospital data. Section 6 discusses the results and our conclusions.

2 Literature review: economies of scope and related measures

The classic definition of economies of scope for multi-output firms stems from Panzar and Willig (1981) and Baumol et al. (1982) that economies of scope exist where it costs less to combine two or more firms with orthogonal output vectors than to continue producing separately. The mathematical definitions for the existence and measurement of economies of scope are expressed through the cost function, $C(y)$, where $M = \{1, \ldots, m\}$ is the set of outputs and $y \in \mathbb{R}^m_+$ is the firm's output vector.

Definition 1 (Baumol et al. 1982) Let $P = \{T_1, \ldots, T_k\}$ denote a non-trivial partition of $S \subseteq M$. There are economies of scope at y_S with respect to the partition P if

$$\sum_{i=1}^{k} C(y_{T_i}) > C(y_S).$$

If the reverse inequality holds, there are diseconomies of scope at y_S. Note that this definition is restrictive in the sense that economies of scope is determined from comparing specialized firms with orthogonal output vectors to a large firm producing the sum of the orthogonal components. Evans and Heckman (1984) observe that these areas of the cost function may not be well characterized when data on these types of firms is unavailable. Revised definitions using less specialized regions of the cost function, and considering other aggregation and dissolution strategies include:

- Berger et al. (1987) define expansion path subadditivity (EPSUB) considering two firms (or two production plans) A and B by

$$EPSUB(y) = \frac{C(y_1^A, y_2^A, y_3^A) + C(y_1^B, y_2^B, y_3^B) - C(y_1^A + y_1^B, y_2^A + y_2^B, y_3^A + y_3^B)}{C(y_1^A + y_1^B, y_2^A + y_2^B, y_3^A + y_3^B)}$$

- Ferrier et al. (1993) define economies of diversification (DIVERS) by

$$DIVERS(y) = \frac{C(y_1^A, 0, y_3^A) + C(0, y_2^B, y_3^B) - C(y_1^A, y_2^B, y_3^A + y_3^B)}{C(y_1^A, y_2^B, y_3^A + y_3^B)}$$

- Preyra and Pink (2006) define a relaxed notion of economies of scope by

$$SCOPE(y) = \frac{C\left(y_1 - \epsilon_1, \frac{\epsilon_2}{m-1}, \cdots, \frac{\epsilon_m}{m-1}\right) + \cdots + C\left(\frac{\epsilon_1}{m-1}, \cdots, \frac{\epsilon_{m-1}}{m-1}, y_m - \epsilon_m\right)}{C(y)}$$

However, these generalizations still suffer the limitations of using price information and entangling scope and scale concepts. Chavas and Kim (2010) address this issue by decomposing economies of diversification into complementarity, scale, convexity, and fixed cost components. Data can be used to estimate the parameters of the cost function when input price data is available and a functional form can be assumed. Once the cost function is parameterized, it is easy to calculate economies of scope, or any of its extensions. For example, Preyra and Pink (2006) estimate a quadratic cost function for hospitals in Ontario, and examine the scope and scale properties of the estimated function to make policy recommendations regarding industry configuration. However, the cost function approach is inappropriate if price data is not available or if firms do not face identical prices. Instead, scope properties may be examined through the production process itself, comparing the production sets of diversified and specialty firms (Färe 1986; Prior 1996; Morita 2003).

3 Returns to scope

Clearly if a firm is a price taker in input and output markets, the only source of economies of scope is through the characteristics of the production function itself, despite its definition and usual estimation through the cost function. Therefore, our approach avoids reliance on prices. Noting that the classic definition of economies of scope and its related extensions are inextricably entangled with the scale properties of the production function, we also suggest that it is more appropriate to define scope properties orthogonal to scale, i.e., based on the trade-offs between outputs at a fixed input level.

We note that output sets may differ depending on the types of inputs used. Consider the classic convex output set illustrated by the case of a mill which uses bundles of wheat to produce both grain and chaff. The solid line in the left panel of Fig. 1 shows the production possibilities characterized by the output set at a fixed level of input. Thus, the output set as determined by a pure material balance argument follows the dashed lines in the left panel. The departure of the production possibilities frontier (the solid line) from the material balance frontier is thought to be greatest when the firm attempts to produce both outputs at a high level because more outputs are lost due to multitasking. For the case of a firm using yarn as an input to produce hats and mittens, the dashed line in the right panel

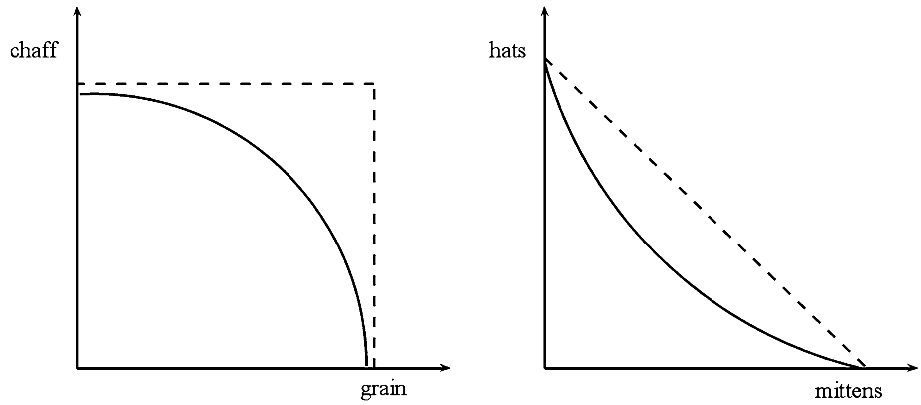

Fig. 1 Material balance and production possibilities

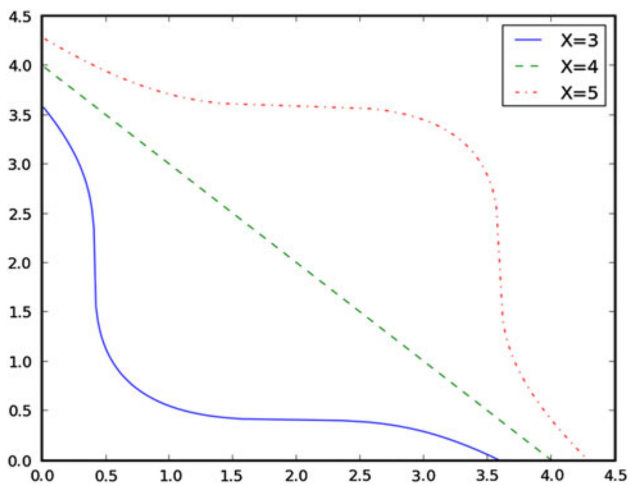

Fig. 2 Varying output frontiers by scale

shows the linear output set determined by the material balance constraint. Similarly, departure from this line in the production possibilities set, potentially created by changeover costs, is greatest when both outputs are relatively high, and produces an output set (the solid line) which is non-convex.

Further, consider the boundary of the output set for a single input, two output production technology. When each output exhibits 'S-shape' production, the shape of the output sets can vary by input level. Let the total input x be split between outputs y_1 and y_2 with proportions p and $(1-p)$. Let each output be produced by $y_j = 2\sqrt[3]{.5x_j - 1} + 2$. Figure 2 shows significant variation in the shape of the boundary by input level.

Another argument used to diminish the need for methods which can characterize non-convex frontiers is that profit maximizing firms will never be observed in these regions of the frontier even if the true frontier is non-convex. Yet, there are several reasons why firms might be observed in such regions, e.g., firms are simply allocatively inefficient and produce in a non-convex region in error. Public firms who make output mix decisions using criteria other than profit maximization could also be observed in non-convex regions, or contractual restrictions can prevent the selection of profit maximizing output mixes. Profit maximizing firms themselves might be observed producing in a non-convex region if the price taking assumption is too strong and a pricing equilibrium supports the firm's behavior. Further, if markets are imperfectly competitive, excess capacity may lead firms to increase their scope of production (Wolinsky 1986). Finally, firms which are regulated, dynamically expanding, or facing uncertainty might also be observed in non-convex regions. For example, risk-averse firms uncertain about production outcomes will avoid specializing in any way which creates high risk. All of these justifications support our interest in characterizing non-convex output sets.

3.1 Convexity in production technologies

Inevitably, scope assessments are influenced by convexity, a popular assumption for production technologies despite the observation that convexity is assumed for its analytical convenience rather than its economic realism (McFadden and Fuss 1978). Several nonparametric estimators such as data envelopment analysis (DEA, Cooper et al. 2006) assume full convexity of the production technology, enforcing convexity of all input sets, output sets, and input–output sets. Of these, the assumption of convexity between inputs and outputs has received the most attention, whereas the assumption of convexity of input sets has received little scrutiny in the literature based on the theoretical justification of diminishing marginal rates of technical substitution (although even this assumption may not hold when the production process could make better use of a single type of input as opposed to a mixture of two types). There has been little interest in relaxing the assumption of convexity of output sets, based on the theory of increasing opportunity costs. The theory of increasing opportunity costs argues that, holding inputs fixed and producing only output 1, the opportunity costs in terms of output 1 required to produce a single unit of output 2 are relatively low. As production of output 2 increases, resources which have higher opportunity costs in terms of output 1 must be used to produce output 2. We note that this argument is equivalent to arguing for increasing absolute marginal rates of transformation between outputs. While popular, these arguments neglect effects such as change-over costs, setup costs, and knowledge spillover, which may vary with scale. For example, as firms increase in scale, complexity and bureaucratic burden may reduce the opportunity for gains from scope (De Witte and Marques 2011). In the case that all resources are equally good at producing either of the two outputs, increasing marginal products or change-over costs would lead to non-convex output sets with a boundary bowed towards the origin. Other support for non-convex output sets appears in Becker and Murphy (1992), who note that "increasing returns from concentrating on a narrower set of tasks raises the productivity of a specialist above that of a jack-of-all-trades." See Briec et al. (2006) for further discussion of convexity assumptions and their relationship to the cost function.

The free disposal hull (FDH) (Afriat 1972) model completely relaxes convexity, leaving only the assumptions of monotonicity, envelopment, and minimum extrapolation. Motivations for relaxing the convexity assumptions could include the desire to accurately estimate productions functions with regions of increasing returns to scale or increasing marginal product and regions where non-convex output sets prevail. The trade-off for additional flexibility in characterizing these regions is diminished ability to discriminate between firms (Wheelock and Wilson 2009). Other researchers have partially relaxed convexity, e.g., Kuosmanen (2001) allows as much convexity as possible while not allowing any FDH efficient firms to appear inefficient. Bogetoft et al. (2000) construct a technology which is convex in both input sets and output sets, but not necessarily so between inputs and outputs.

3.2 Definition and relationship to economies of scale

Before formalizing a definition of returns to scope, recall the arguments which have been presented regarding plausible trade-offs between outputs in multi-output production theory (Baumol et al. 1982, Section 4C). One argument arising from the work of Adam Smith is that returns to specialization could exist between outputs. Explanations for such returns include the notion that laborers can produce a single output more efficiently than multiple outputs because of learning, repetition, or change over costs. An alternative argument comes from the Marshallian notion of joint production, and suggests that a firm diversified in its outputs should be able to produce more efficiently than two specialized firms. This argument relates to the notion of quasi-public inputs. Turning again to Fig. 1, the left panel displays Marshallian returns to diversification and convex output sets, while the right panel illustrates Smith's concept of returns to specialization with non-convex output frontiers bowed towards the origin.

Our definition of returns to scope uses these concepts and theoretical justifications for investigating the shape of the output isoquant. Let $T = \{(x, y) | x \text{ can produce } y\} \subseteq \mathbb{R}_{+}^{n+m}$ represent the production technology, which can be equivalently expressed through its input sets, $I(y) = \{x | (x, y) \in T\}$, or output sets, $P(x) = \{y | (x, y) \in T\}$. A production technology is said to exhibit constant returns to scale (CRS) if $(x, y) \in T \Rightarrow (\lambda x, \lambda y) \in T, \forall \lambda \geq 0$. Let $B(P(x))$ be the boundary of the output set $P(x)$. Our notion of returns to scope is defined between pairs of outputs, with inputs and other outputs fixed, so we frequently use the notation $P(x, y_{-ij}) = \{(y_i, y_j) | (x, y_i, y_j, y_{-ij}) \in T\}$. Since only non-negative quantities of inputs and outputs are meaningful, all set complements are restricted to the non-negative orthant, i.e., $P^c(x) = \{y \in \mathbb{R}_{+}^{m} | (x, y) \notin T\}$. We make two standard assumptions regarding the production technology:

Assumption 1: T is monotonic, that is $(x, y) \in T \Rightarrow (x', y') \in T \; \forall x' \geq x, y' \leq y$.

Assumption 2: T is compact in output sets, i.e., $\forall x, P(x)$ is compact.

Definition 2 (i) Positive returns to scope exist between outputs i and j at x, y_{ij} if $P(x, y_{-ij})$ is convex.

(ii) Negative returns to scope exist between outputs i and j at x, y_{ij} if $P^c(x, y_{-ij})$ is convex.

If the production technology T can be characterized by the multi-input, multi-output production function $F(x, y) \geq 0$, with frontier $F(x, y) = 0$, and satisfying the conditions of the implicit function theorem, the following equivalent definition can be stated.

Definition 3 (i) Positive returns to scope exist between y_i and y_j at x, y_{-ij} *if* $\frac{\partial^2 y_i}{\partial^2 y_j} \leq 0$ along $B(P(x, y_{-ij}))$.

(ii) Negative returns to scope exist between y_i and y_j at x, y_{-ij} *if* $\frac{\partial^2 y_i}{\partial^2 y_j} \geq 0$ along $B(P(x, y_{-ij}))$.

If both positive and negative returns to scope exist, outputs i and j have a linear returns to scope at x, y_{ij}. Based on our definition of returns to scope, we consider the relationship between returns to scope, returns to scale, and economies of scope.

Proposition 1 *For CRS technologies, the assessment of returns to scope between two outputs is constant along all linear expansion and contraction paths.*

Proof: We prove the consistency of a positive returns to scope assessment between two outputs, using Definition 2, from which the proofs for negative and linear returns are analogous. Suppose T exhibits CRS and positive returns to scope exist between outputs i and j at x', y'_{-ij}. Then we need to ensure the convexity of $P(\lambda x, \lambda y_{-ij})$ for $\lambda > 0$. Let $y^1 = \left(y_i^1, y_j^1\right), y^2 = \left(y_i^2, y_j^2\right) \in P(\lambda x, \lambda y_{-ij})$ for $\lambda > 0$. Then by CRS, $\frac{y^1}{\lambda}, \frac{y^2}{\lambda} \in P(x, y_{-ij})$. Positive returns to scale implies that $\frac{\theta y^1}{\lambda} + \frac{(1-\theta)y^2}{\lambda} \in P\left(x, y_{-ij}\right), \forall \theta \in [0, 1]$, and CRS again implies that $\theta y^1 + (1-\theta)y^2 \in P(\lambda x, \lambda y_{-ij}), \forall \theta \in [0, 1]$. □

We also state the following result regarding the relationship between returns to scope and economies of scope.

Proposition 2 *Concerning single-input, two-output production technologies:*

(i) Technologies which are characterized by positive returns to scope at all input levels and non-decreasing returns to scale display economies of scope for all pairs of production plans.

(ii) Technologies which are characterized by negative returns to scope at all input levels and non-increasing returns to scale display diseconomies of scope for all production plans.

Proof: (i) Consider two production plans $(y_1, 0)$ and $(0, y_2)$. In order to observe economies of scope, it must be that $C(y_1, y_2) < C(y_1, 0) + C(0, y_2)$. A sufficient condition for this cost relation to be observed is that $L(y_1, 0) + L(0, y_2) \subseteq L(y_1, y_2)$. Let x_1 be the minimum amount of input that can produce $(y_1, 0)$ and x_2 be the minimum amount of input that can produce $(0, y_2)$. Then $L(y_1, 0) = \{x : x \geq x_1\}, L(0, y_2) = \{x : x \geq x_2\}$, and $L(y_1, 0) + L(0, y_2) = \{x : x \geq x_1 + x_2\}$. All that is left to show is that $x_1 + x_2$ can produce (y_1, y_2).

Without loss of generality, let $x_2 \geq x_1$, and let $\frac{x_2}{x_1} = \lambda \geq 1$. Then by non-decreasing returns to scale, $(1+\lambda)(y_1, 0) \in P(x_1 + x_2)$ and $\left(1 + \frac{1}{\lambda}\right)(0, y_2) \in P(x_1 + x_2)$. Now by positive returns to scope, and choosing $\theta = \frac{1}{1+\lambda}, \exists y \in P(x_1 + x_2)$ such that $y \geq \theta(1+\lambda)(y_1, 0) + (1-\theta)\left(1 + \frac{1}{\lambda}\right)(0, y_2) = (y_1, y_2)$. Then monotonicity implies $(x_1 + x_2) \in L(y_1, y_2)$.

(ii) Consider a production plan (y_1, y_2). In order to observe diseconomies of scope, it must be that $C(y_1, y_2) > C(y_1, 0) + C(0, y_2)$. A sufficient condition is that $L(y_1, 0) + L(0, y_2) \supseteq L(y_1, y_2)$. Let x_{12} be the minimum amount of input that can produce (y_1, y_2). Let $(\widehat{y}_1, 0)$ be the maximum amount of the first output that x_{12} can produce. By our assumptions of NRSc, the frontier cannot be bowed away from the origin more than linearly at any point and combined with our assumption of compact output sets, this forces $\widehat{y}_1 > y_1$. Again, by NRSc the frontier is bowed towards the origin at least linearly at every point, so the maximum amount of the second output producible by x_{12} is at least $\widehat{y}_2 = \frac{y_2}{\widehat{y}_1 - y_1}\widehat{y}_1 > y_2$. Now let $\lambda_1 = \frac{y_1}{\widehat{y}_1} < 1$, and $\lambda_2 = \frac{y_2}{\widehat{y}_2} < 1$. By our assumption of non-increasing returns to scale, $\lambda_1 x_{12}$ can produce y_1, and $\lambda_2 x_{12}$ can produce y_2. Finally, $\lambda_1 + \lambda_2 = 1$, so $x_{12} \in L(y_1, 0) + L(0, y_2)$, and diseconomies of scope must be observed. □

Proposition 2 shows that returns to scope and economies of scope are related, but non-nested. Even simple technologies (single-input, two-output) exist which exhibit positive returns to scope at all input levels, yet exhibit diseconomies of scope. Similarly, there could exist technologies which exhibit negative returns to scope at all input levels, but economies of scope could be found. We illustrate the potentially misleading signal from economies of scope using a single-input, two-output production function $x^\beta = y_1^\alpha + y_2^\alpha$. We can create examples of production function which exhibit negative (positive) returns to scope everywhere and have economies (diseconomies) of scope. Graphs of these two cases appear in Fig. 3. First let $\alpha = .5, \beta = 1.5$, this production function displays negative returns to scope everywhere. If we normalize the cost per unit of x to 1, we can quickly compute that $C(5, 0) + C(0, 5) = 3.43 > 2.73 = C(5, 5)$, finding economies of scope. Now let $\alpha = 2, \beta = .5$, this production function displays positive returns to scope everywhere. We compute that $C(5, 0) + C(0, 5) = 1250 < 2500 = C(5, 5)$, finding diseconomies of scope.

Capturing the inverse relationship, what economies of scope imply about returns to scope, is elusive due to the composite nature of the economies of scope measure and the flexibility of returns to scope to vary between input levels. We define returns to scope with this in mind, since it is plausible for small firms to exhibit negative returns to scope while larger firms exhibit positive returns to scope.

From a policy perspective, scope properties help to answer industry configuration questions, such as, "Given

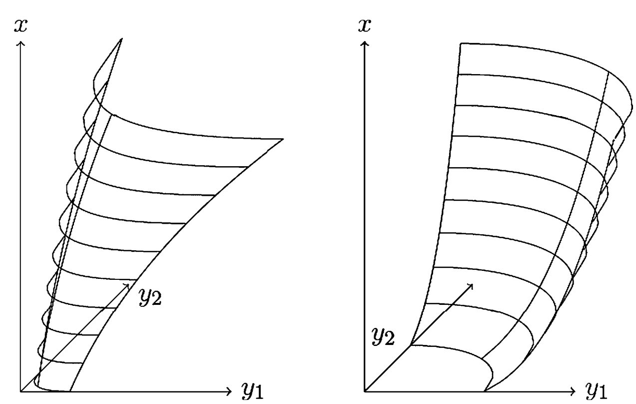

Fig. 3 Returns to scope, economies of scope, and scale of production

two hospitals of the same size (input level), should they specialize or diversify between outputs, e.g., diagnostic and therapeutic procedures, to achieve the greatest levels of productivity?" Using the new returns to scope measure, the answer is diversify if the production process exhibits positive returns to scope at the input level, and specialize if the production technology exhibits negative returns to scope at the input level. Whether these hospitals currently exhibit economies or diseconomies of scope is irrelevant. The knowledge of returns to scope provides information for policy-makers in structuring industries, managers in production planning, and analysts in modeling and benchmarking. We now introduce two methods for assessing returns to scope.

4 Methods for assessing returns to scope

From the existing literature, we identify five properties desirable for our assessment method:

1. Applicability without price data or observations of specialized firms.
2. Ability to detect both positive and negative returns to scope.
3. Avoid unnecessarily strong assumptions about the functional form of the production function.
4. Robust to noise in the data.
5. Use of frontier data.

Below we present two methods for identifying returns to scope.

4.1 Method one: output frontier regression

This two-stage method is similar to Thiry and Tulkens (1992) and Bardhan et al. (1998), who report improved results from first identifying efficient observations before performing estimation. The first stage identifies the efficient subset of the data. Since scope and scale are properties of the production frontier, they are appropriately estimated using only frontier observations. In order to identify efficient firms, Assumptions 1 and 2 from Sect. 3.2 are used. Recall that under these assumptions the efficient observations at any input level x are those which use at most x input and are non-dominated in outputs. Observation k with production vector (x^k, y^k) is efficient at input level x if $x^k \leq x$ and $\nexists$ another observation, say l with $x^l \leq x$ and $y^l \geq y^k$. This criterion selects the efficient observations from the FDH technology at each input level. The second stage fits a curve for each pair of outputs by using the efficient points at a given level of input to estimate returns to scope between a pair of outputs. The curve we estimate by OLS is

$$0 = F(x, y) = x + \alpha_i y_i + \alpha_j y_j + \beta_{ij} y_i y_j,$$

which holding x constant gives

$$\frac{\partial y_j}{\partial y_i} = -\frac{\partial F}{\partial y_i} \Big/ \frac{\partial F}{\partial y_j},$$

$$\frac{\partial^2 y_j}{\partial y_i^2} = -\frac{\frac{\partial y_j}{\partial y_i} \frac{\partial^2 F}{\partial y_i \partial y_j}}{\frac{\partial F}{\partial y_j}}.$$

Monotonicity ensures that the sign of $\frac{\partial^2 y_j}{\partial y_i^2}$ is constant, and thus we can use Definition 3 to assess the returns to scope. Since $\frac{\partial^2 F}{\partial y_i \partial y_j} = \beta_{ij}$, the sign of the β_{ij} term gives the assessment of returns to scope. That is, a negative value for β_{ij} implies that $\frac{\partial^2 y_j}{\partial y_i^2}$ is positive, which we defined as negative returns to scope. In the case that the production technology uses multiple inputs, a measure of aggregate input, $g(x)$, can replace x.

An appealing feature of this method is the statistical significance of the β_{ij} parameter provides certainty about whether the data supports a conclusion of positive or negative returns to scope. Although the first stage is sensitive to noise in the data, we find that this sensitivity can be reduced by using a thick frontier method, e.g., Berger and Humphrey (1991). The functional form of the pairwise frontier chosen is appealing because of its simplicity, ease of estimation, and flexibility to characterize positive or negative returns to scope.

4.2 Method two: comparison of hull areas

This method compares the areas of hulls constructed using various estimation procedures. DEA estimates convex output sets which are present when the production process features positive returns to scope. FDH estimates non-convex output sets, more closely approximating the production frontier when non-convexities are present. We note that the areas from DEA and FDH efficient points should not be directly compared since in the case that convexity does hold, these procedures will likely identify the same set of efficient points, but yield different areas of output sets. Instead we need to construct a baseline to which we can compare a DEA output set area to identify positive returns to scope, and to compare an FDH output set area to identify negative returns to scope. We use the definition of anchor points from Bougnol and Dulá (2009) to construct the anchor interpolation hull (AIH) for this purpose.

Definition 4 (Bougnol and Dulá 2009) An extreme-efficient point is an anchor point if it belongs to an unbounded face of $\mathcal{P}^{VRS}$.

The set of extreme-efficient points is the minimal set of points used to define a convex hull which is equivalent to the variable returns to scale (VRS) DEA hull $\mathcal{P}^{VRS}$. Bougnol and Dulá prove that an extreme-efficient point is an anchor if and only if there exists a solution to the DEA multiplier formulation where one of the multiplier prices is zero. They also give an algorithm for identifying anchor points. We construct the anchor interpolation hull (AIH) from anchor points, and compare it to hulls defined by DEA and FDH efficient points.

To assess returns to scope between outputs i and j at input level x, this method uses all data points with inputs less than x. We map the outputs into $\mathbb{R}^2$ by the ij-projection mapping, $\pi_{ij}(y_1,\ldots,y_m) = (y_i, y_j)$. An alternative to using projections would be to use sections such as those described in Krivonozhko et al. (2004), however these sections assume convexity. The three areas used to assess returns to scope are based on convex polytopes

$$A_{DEA} = A\left(conv\left(Y^{eff}_{DEA} \cup \mathbf{0}\right)\right), \tag{1}$$

$$A_{FDH} = A\left(conv\left(Y^{eff}_{FDH} \cup \mathbf{0}\right)\right) - A\left(conv(Y^{eff}_{FDH})\right), \tag{2}$$

$$A_{AIH} = A\left(conv\left(Y^{eff}_{anchor} \cup \mathbf{0}\right)\right), \tag{3}$$

where $conv(S)$ means the convex hull of the set S, Y^{eff}_{model} represents the set of points deemed as efficient using a particular model, and $A(\cdot)$ is the area operator. Note that since anchor points are always DEA efficient, and DEA efficient points are always FDH efficient, meaning all three sets of efficient points include the anchor points. Therefore, the only difference in areas comes from the shape of the frontier characterizing the trade-off between outputs i and j. The ranking of the areas is given by $A_{FDH} \le A_{AIH} \le A_{DEA}$. This leads to the following definition:

Definition 5 For a given input level x, and considering outputs i and j, let A_{DEA}, A_{FDH}, and A_{AIH} be the areas given by (1)–(3). Then we have the following measures for positive returns to scope (PRSc) and negative returns to scope (NRSc):

$$PRSc = \frac{A_{DEA} - A_{AIH}}{A_{AIH}}, \tag{4}$$

$$NRSc = \frac{A_{AIH} - A_{FDH}}{A_{AIH}}. \tag{5}$$

By construction, both $PRSc$ and $NRSc$ are greater than or equal to 0. In general, both $PRSc$ and $NRSc$ could be greater than 0. The production process exhibits positive returns to scope between i and j at x if $\max\{\text{PRSc}, \text{NRSc}\} = PRSc$, and negative returns to scope if $\max\{\text{PRSc}, \text{NRSc}\} = NRSc$. The area-based approach is more sensitive to noise in the data than the two-stage method, but in the case that noise is bounded and uniformly affects all mixes of outputs, Gstach (1998) shows that envelopment frontiers are asymptotically consistent estimators for the shape of the frontier. Noise can be bounded by using outlier detection procedures (Johnson and McGinnis 2008). The advantage of the area-based method is the lack of a functional form specification.

Note that our definitions of returns to scope are for pairs of outputs. For firms with three or more outputs, both of our methods assess the pairwise trade-offs between y_i and y_j independent of the level of y_k. We make an additional assumption that returns to scope between y_i and y_j do not vary with y_k, primarily to reduce the amount of data required for assessment of returns to scope. By our assumption of monotonicity, the ij-output set will clearly vary with y_k, and our implicit assumption is only that the shape of the frontier does not vary with y_k. We call this property output scope invariance, a relaxation of the property that ij-output sets are homothetic in y_k. We discuss how this assumption can be tested in the "Appendix".

5 Illustrative example and hospital application

To test the effectiveness of these methods, we first consider a single-input, two-output production function. The input, x, is drawn uniformly from the integers $\{5, 6, 7, \ldots, 15\}$. Inefficiency, u, is generated from $|N(0, 0.1)|$, and enters the production process as wasted input. Using total resources $(1-u)x$, a firm then devotes a fraction, $p \sim U(0,1)$, of these resources to produce the first output, y_1, and the remaining resources, $(1-p)(1-u)x$, to produce the second output, y_2. Based on resources r_i used in producing y_i, the amount of output generated is given by $y_i = r_i^{1/\gamma}$. We vary $\gamma \in \{0.5, 1, 2\}$ to create output set frontiers bowed towards the origin, displaying negative returns to scope; linear; and bowed away from the origin, displaying positive returns to scope. We experiment with 50 and 100 firms with the results reported in Tables 1 and 2. Statistical significance of method one is indicated by *, **, and *** for significance levels 0.10, 0.05, and 0.01. Although method two has no statistical basis for significance, we highlight large values of $|PRSc - NRSc| > 0.15, 0.30,$ and 0.45, by ^, ^^, ^^^.

Whether using 50 or 100 firms, the two methods agree in their assessment of returns to scope at all input levels when $\gamma = 0.5$ or 2. The sign of the β_{12} term is significant for many of these tests and the magnitude of $|PRSc - NRSc|$ is large, especially in the runs with 100 firms. When $\gamma = 1$, the output set frontier is linear, and both methods show little evidence for positive or negative returns to scope. The two-stage method finds no significant returns to scope at

Table 1 Example of methods for assessing returns to scope: 50 firms

$\gamma = 0.5$			$\gamma = 1$			$\gamma = 2$		
x	$\beta 12$	PRSc-NRSc	x	$\beta 12$	PRSc-NRSc	x	$\beta 12$	PRSc-NRSc
5	−0.043	−0.430^^	5	0.076	0.050	5	0.153	0.046
6	−0.054	−0.434^^	6	−0.015	−0.144	6	0.536	0.022
7	−0.017	−0.626^^	7	−0.049	−0.101	7	−0.025	0.060
8	−0.010***	−0.369^^	8	−0.041	0.043	8	1.112	0.097
9	−0.011***	−0.466^^^	9	0.008	0.173^	9	1.096	0.154^
10	−0.015***	−0.678^^^	10	0.067	0.159^	10	−0.043	0.075
11	−0.013***	−0.734^^^	11	0.056	−0.031	11	0.613	0.105
12	−0.008**	−0.744^^^	12	0.019	−0.219^	12	0.966**	0.207^
13	−0.005**	−0.560^^^	13	0.023	−0.300^^	13	1.224	0.117
14	−0.002*	−0.284^	14	0.006	0.096	14	1.647***	0.100
15	−0.002**	−0.642^^^	15	0.007	−0.019	15	0.826*	0.159^

Table 2 Example of methods for assessing returns to scope: 100 firms

$\gamma = 0.5$			$\gamma = 1$			$\gamma = 2$		
x	$\beta 12$	PRSc-NRSc	x	$\beta 12$	PRSc-NRSc	x	$\beta 12$	PRSc-NRSc
5	−0.056***	−0.473^^^	5	0.002	−0.035	5	0.952**	0.222^
6	−0.074***	−0.707^^^	6	−0.014	−0.226^	6	0.913**	0.330^^
7	−0.021***	−0.446^^	7	−0.016	−0.119	7	0.716	0.191^
8	−0.020***	−0.774^^^	8	0.014	0.117	8	0.810***	0.221^
9	−0.014***	−0.743^^^	9	−0.026	−0.191^	9	0.774**	0.257^
10	−0.012***	−0.722^^^	10	0.016	−0.015	10	0.985**	0.324^^
11	−0.007***	−0.585^^^	11	0.005	−0.101	11	0.806**	0.115
12	−0.004***	−0.559^^^	12	0.009	0.055	12	0.745***	0.227^
13	−0.003***	−0.492^^^	13	0.000	−0.047	13	0.745***	0.324^^
14	−0.003***	−0.598^^^	14	0.009	0.133	14	0.669***	0.228^
15	−0.003***	−0.681^^^	15	0.012	−0.269^	15	0.781***	0.276^

any input level for either experiment. Both methods appear to benefit from more data. The two-stage method finds more significance from 100 firms when $\gamma = 0.5$ or 2. The area-based method shows greater differences in PRSc and NRSc when $\gamma = 0.5$ or 2, and smaller differences when $\gamma = 1$.

For data with many distinct input levels, the range of inputs can be discretized to assess returns to scope. In general, returns to scope are potentially different at each input level. However, for certain technologies, e.g., output homothetic technologies (Färe and Primont 1995), scope assessments should be consistent at all input levels. In the case that the technology is assumed to be output-homothetic, our methods can test this assumption.

To investigate the robustness of our results, we run 100 trials at each γ level for each experiment size. Table 3 reports the proportion of significant (and pseudo-significant) positive or negative returns to scope by each of the two methods. Checking our assumption of monotonicity from the regression results, we find that monotonicity holds at 100 % of FDH efficient firms for more than 90 % of all experiments. Table 3 also shows that at $\gamma = 1$ method one may be biased towards PRSc and method two may be biased towards NRSc. Despite creating bias in small samples, when output sets are convex (such as when $\gamma = 1$), we know that both DEA and FDH are asymptotically consistent estimators of the true frontier (Banker 1993; Korostelev et al. 1995). Therefore as the number of firms becomes large, the set of DEA and FDH efficient points will coincide in the limit and the biases of each method will vanish. The source of each methods bias is discussed in the

Table 3 Portion of significant (and pseudo-significant) trials for two-output examples

γ	50 firms				100 firms			
	Two-stage		Area		Two-stage		Area	
	PRSc	NRSc	PRSc	NRSc	PRSc	NRSc	PRSc	NRSc
0.5	0.001	0.756	0.003	0.970	0.000	0.961	0.000	0.998
1	0.074	0.030	0.101	0.333	0.071	0.031	0.080	0.241
2	0.580	0.002	0.541	0.002	0.913	0.001	0.761	0.000

"Appendix". Practically speaking, these biases can be useful, for example when the number of firms is not infinite, the tendency of each method to report opposite conclusions when output frontiers are linear can be used as a robustness check. When both methods report the same returns to scope, we have strong evidence to believe the results. When they differ, further investigation may be warranted.

Experiments with 3, 4, and 5 output firms show that both methods perform well in higher dimensions. To simulate firms with k-dimensional outputs we draw a random vector from the k-simplex to simulate the fraction of the resources the firm directs to producing each output. We implement random sampling from the k-simplex by first generating k random variates $z_i \sim expo(1)$, and then dividing each of the k random variates by the sum their sum $\sum z_i$. We compute the monotonicity constrained coefficients via quadratic programming and find the differences from the unconstrained regression coefficients are small, although monotonicity is more frequently violated in higher dimensions. Further results are available from the authors upon request.

6 Application: hospitals as multi-output firms

Many studies have modeled hospitals as multi-output firms (see Hollingsworth 2008; Rosko and Mutter 2011 for reviews) including several studies which consider scope properties (Preyra and Pink 2006; Smet 2007; Ferrier et al. 2009). In order to examine returns to scope, we use the AHRQ HCUP 2008 Nationwide Inpatient Sample, a data set which contains all discharges from an approximate 20 % sample (1,056 hospitals) of US community hospitals as defined by the American Hospital Association. The sample is stratified across hospital bed sizes (small/medium/large), location (urban/rural), control type, geographical region, and teaching status. We model 4 outputs: minor diagnostic procedures (y_1), major diagnostic procedures (y_2), minor therapeutic procedures (y_3), and major therapeutic procedures (y_4), categorized by International Classification of Diseases, Clinical Modification (ICD-9-CM) codes. The distinguishing characteristic between minor and major procedures of each type is the use of an operating room. For example, a CT scan is a minor diagnostic procedure, whereas a brain biopsy is a major diagnostic procedure, an irrigate ventricular shunt is a minor therapeutic procedure, whereas an aorta-renal bypass is a major therapeutic procedure. We hypothesize that small hospitals exhibit negative returns to scope, and larger hospitals positive returns to scope. This hypothesis is based on the idea that resources in smaller hospitals will be allocated to producing a variety of outputs, whereas resources in larger hospitals can be allocated to specialized tasks. Using number of discharges as a single input proxy for the hospital's aggregate input level (x), we test for returns to scope between outputs at 9 input levels. Table 4 shows the input levels, the number of hospitals in each input level 'bin', and the number of FDH efficient hospitals at each input level. Tables 5 and 6 report the results of the two assessment methods.

Table 7 combines the reports of returns to scope. Cells in which the tests agree or only one test gives conclusive evidence are marked according to the returns to scope found. Cells where the tests disagree or both tests give weak evidence are left blank.

Table 4 Hospital counts by input levels

k	1	2	3	4	5	6	7	8	9
$\bar{x}_k$	250	500	1,000	2,000	4,000	8,000	16,000	32,000	64,000
$\#\{i : x_i \in (\bar{x}_{k-1}, \bar{x}_k]\}$	62	102	117	136	140	156	167	140	25
$\#\{i : i \in FDH(\bar{x}_k)\}$	11	18	11	13	11	11	23	8	4

Table 5 Two-stage method assessment of hospital returns to scope

$\bar{x}$	y_1y_2	y_1y_3	y_1y_4	y_2y_3	y_2y_4	y_3y_4
250	−0.472	0.001	−0.015	0.042	1.964*	0.002
500	0.090	0.002***	0.002	−0.009	0.140	−0.000
1,000	0.160**	0.002*	−0.001	0.053***	0.009	0.000
2,000	0.015	0.001*	0.000	0.016***	0.003	−0.000
4,000	0.007	0.000	0.000	0.005	0.005	0.000
8,000	0.003	0.000***	0.000**	0.002	0.011**	0.000***
16,000	0.004***	0.000***	0.000***	0.002*	0.006***	0.000***
32,000	0.004**	0.000***	0.000*	−0.000	0.001*	0.000*
64,000	0.000	0.000**	−0.000	0.000	0.001	0.000

Table 6 Area-based method assessment of hospital returns to scope

$\bar{x}$	y_1y_2	y_1y_3	y_1y_4	y_2y_3	y_2y_4	y_3y_4
250	−0.817^^^	−0.255^	−0.555^^^	−0.544^^^	−0.450^^	−0.018
500	−0.152^	−0.192^	0.037	−0.615^^^	0.109	−0.520^^^
1,000	0.000	0.112	−0.529^^^	−0.126	0.072	−0.367^^
2,000	−0.551^^^	0.115	−0.218^	0.000	−0.261^	−0.293^
4,000	−0.204^	−0.051	−0.153^	0.000	0.035	−0.217^
8,000	−0.228^	0.000	0.000	0.046	0.092	0.000
16,000	−0.290^	0.117	−0.197^	−0.226^	−0.188^	0.000
32,000	−0.169^	0.000	0.000	0.037	0.135	−0.149
64,000	0.000	0.000	0.064	−0.027	0.000	0.000

Table 7 Combined assessment of hospital returns to scope

$\bar{x}$	y_1y_2	y_1y_3	y_1y_4	y_2y_3	y_2y_4	y_3y_4
250	−	−	−	−		
500	−			−		−
1,001	+	+	−	+		−
2,000	−	+	−	+	−	−
4,000	−		−			−
BOQO	−	+	+		+	+
16,000		+				+
32,000	−	+	+		+	+
64,000		+				

We gain three important insights from the analysis. First, at 6 of 9 input levels there is evidence of negative returns to scope between minor and major diagnostic procedures (y_1, y_2). The contrasting resources needed for these outputs, e.g., major diagnostic procedures require highly specialized instrumentation and labor, whereas minor procedures require less, make it plausible that hospitals producing both minor and major diagnostics will spend considerable time switching between the two types of procedures and thus produce less than hospitals specializing in either output. The only evidence to the contrary appears at input level 1,000, which showed positive returns to scope. However, our investigation revealed that a single hospital at this input level produced 45 major diagnostic procedures (no other FDH efficient hospital at this input level produced more than 13) and produced 609 minor diagnostic procedures (the second most of any hospital at this input level). This observation forces both area-based measures of returns to scope to zero and pulls the two-stage measure of returns to scope positive. Removing this hospital from the data set produced an insignificant measure of returns to scope from the two-stage method, and changed the area-based measure of returns to scope from 0 to −0.547. Second, there is evidence of positive returns to scope at 6 of 9 input levels between minor diagnostic and minor therapeutic procedures (y_1, y_3). Intuitively, these outputs require the same resources, e.g., nursing labor and exam rooms, and therefore share setup costs of production, which leads to positive returns to scope between these outputs. Third, the evidence supports our hypothesis that smaller hospitals display negative returns to scope between all pairs of outputs. The negative returns to scope disappear as hospitals become larger and better suited to produce multiple outputs. A possible explanation for this insight is that larger hospitals are able to share generic resources across similar outputs, thus improving utilization, whereas smaller hospitals have to divide their resources in order to produce distinct outputs. Switching, change-over, or coordination costs are the type of phenomenon we expect to create output sets bowed towards the origin, displaying negative returns to scope. Our findings of negative returns to scope partially echo the results of Smet (2007) who finds positive pair-wise cost complementarities between days of internal medicine, surgery, and specialized care produced in Belgian hospitals.

7 Conclusions

Scope properties can assist strategic decision-making related to multi-output production processes. Extending the extant knowledge of such properties, this paper introduced a new property, returns to scope, which is orthogonal to the scale of production and does not rely on price information. We discussed how returns to scope and the more traditional economies of scope are non-nested concepts, useful for answering different questions about industry structure and product mix strategies. Knowledge of returns to scope also provides information about the shape of the production frontier, and is valuable in determining the most appropriate frontier estimation procedures.

Based on the definition of returns to scope, we present two methods for assessment, which are orthogonal to scale, can detect positive or negative returns to scope, and required only input and output data. Using simulated data we give examples of the ability of each method to detect returns to scope. Although each method possesses a small sample bias, we identify the causes and show that the biases vanished asymptotically. The biases also are in opposite directions, meaning that the assessments are robust when the methods agree.

We use the AHRQ HCUP 2008 National Inpatient Sample to assess returns to scope in hospital production of minor and major diagnostic and therapeutic. Our results illuminate three insights regarding the trade-offs between services provided in hospitals. First, hospitals of all sizes experience negative returns to scope between minor and major diagnostic procedures, implying that when possible these two service lines should be diversified. Second, hospitals of all sizes experience positive returns to scope between minor diagnostic and therapeutic procedures, implying that when possible these two service lines should be jointly produced. Third, small hospitals generally experience negative returns to scope, whereas larger hospitals more often realize positive returns to scope. Based on our findings that scope properties can fluctuate throughout regions of production, functional estimators which are able to selectively impose returns to scope properties are desirable. We leave development of such estimators for future research.

Acknowledgments This research has been funded in part by NSF grant CMMI-0944670. The first author wishes to acknowledge financial support for this work by the US Department of Education through GAANN P200A060241.

Appendix

Testing for output scope invariance

The methods we present for testing for returns to scope ignore the possibility that variation in a third output (y_k) may affect the returns to scope between y_i and y_j. We referred to this property of a production function as output scope invariance. In order to test for the effects of other outputs on the returns to scope between y_i and y_j, the two-stage method can be augmented to test for three way interactions among outputs,

$$0 = F(x, y) = x + \alpha_i y_i + \alpha_j y_j + \beta_{ij} y_i y_j + \sum_{k \neq i,j} \rho_{ijk} y_i y_j y_k.$$

In the case that the ρ_{ijk} terms are insignificant, there is evidence that outputs y_i and y_j are output scope invariant of output y_k.

Small sample biases in assessment methods

The biases of each method can best be explained by considering the data generation process and the details of each method. Figure 4 shows a plot of firms who have at most x' level of input. When $\gamma = 1$, the frontier is linear, and without noise, efficient firms with input level of exactly x' lie farthest from the origin on a line. These points are marked by filled circle. Moreover, the FDH frontier could contain firms with lower input levels, or firms which produce less output due to inefficiency, and would appear between efficient firms at input level x', and are marked by filled diamond. The location of these types of firms explains both the negative bias of the area-based method and the positive bias of the two-stage method. Using this data, the area-based method finds no evidence of PRSc, but using the convex hull of FDH efficient points (the region enclosed by the dashed line in Fig. 4), it finds evidence for NRSc. The two-stage method would estimate a regression line similar to the one shown as a solid line in Fig. 4. If the weakly efficient firms were not part of the sample, the regression line would run directly through all the efficient points, producing $\alpha_1 = \alpha_2 = -1$ and $\beta_{12} = 0$. Adding the presence of the weakly efficient firms has two effects on the estimated parameters. In order to minimize the sum of squared errors, the OLS procedure increases the magnitude of the α_i parameters, making them more negative. To balance this effect, OLS also slightly increases the β_{12} parameter, making it positive, and giving the impression of positive returns to scope.

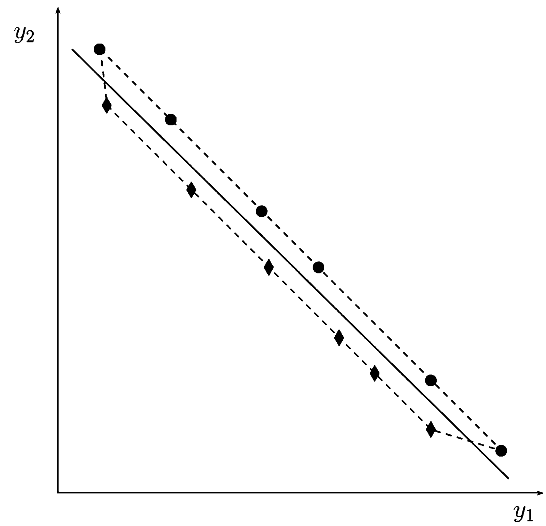

Fig. 4 Small sample biases in returns to scope methods

References

Afriat S (1972) Efficiency estimation of production functions. Int Econ Rev 13:568–598

Arora A, Gambardella A, Magazzini L, Pammolli F (2009) A breath of fresh air? Firm type, scale, scope, and selection effects in drug development. Manage Sci 55(10):1638–1653

Banker RD (1993) Maximum-likelihood, consistency and data envelopment analysis—a statistical foundation. Manage Sci 39(10):1265–1273

Bardhan IR, Cooper WW, Kumbhakar SC (1998) A simulation study of joint uses of data envelopment analysis and statistical regressions for production function estimation and efficiency evaluation. J Prod Anal 9(3):249–278

Baumol WJ, Panzar JC, Willig RD (1982) Contestable markets and the theory of industry structure. Harcourt Brace Jovanovich, Inc., New York

Becker GS, Murphy KM (1992) The division-of-labor, coordination costs and knowledge. Q J Econ 107(4):1137–1160

Berger AN, Humphrey DB (1991) The dominance of inefficiencies over scale and product mix economics in banking. J Monet Econ 28(1):117–148

Berger AN, Hanweck GA, Humphrey DB (1987) Competitive viability in banking—scale, scope, and product mix economies. J Monet Econ 20(3):501–520

Bogetoft P, Tama JM, Tind J (2000) Convex input and output projections of nonconvex production possibility sets. Manage Sci 46(6):858–869

Bougnol M-L, Dulá JH (2009) Anchor points in DEA. Eur J Oper Res 192(2):668–676

Briec W, Kerstens K, Eeckaut PV (2006) Non-convex technologies and cost functions: definitions, duality and nonparametric tests of convexity. J Econ 87(3):307–308

Chavas J-P, Kim K (2010) Economies of diversification: a generalization and decomposition of economies of scope. Int J Prod Econ 126(2):229–235

Cooper WW, Seiford LM, Tone K (2006) Data envelopment analysis: a comprehensive text with models, applications, references and DEA-Solver Software, 2nd edn. Springer, Berlin

De Witte K, Marques RC (2011) Big and beautiful? On non-parametrically measuring scale economies in non-convex technologies. J Prod Anal 35(3):213–226

Evans DS, Heckman JJ (1984) A test for subadditivity of the cost function with an application to the Bell System. Am Econ Rev 74(4):615–623

Färe R (1986) Addition and efficiency. Q J Econ 101(4):861–865

Färe R, Primont D (1995) Multi-output production and duality: theory and applications. Springer, Berlin

Ferrier GD, Grosskopf S, Hayes KJ, Yaisawarng S (1993) Economies of diversification in the banking industry—a frontier approach. J Monet Econ 31(2):229–249

Ferrier G, Leleu H, Moises J, Valdmanis V (2009) The size and service offering efficiencies of US Hospitals. Working paper

Growitsch C, Wetzel H (2009) Testing for economies of scope in european railways an efficiency analysis. J Transp Econ Policy 43:1–24

Gstach D (1998) Another approach to data envelopment analysis in noisy environments: DEA+. J Prod Anal 9(2):161–176

Henderson R, Cockburn I (1996) Scale, scope, and spillovers: the determinants of research productivity in drug discovery. RAND J Econ 27(1):32–59

Hollingsworth B (2008) Efficiency measurement in health and health care. Routledge, London

Johnson AL, McGinnis LF (2008) Outlier detection in two-stage semiparametric DEA models. Eur J Oper Res 187(2):629–635

Korostelev AP, Simar L, Tsybakov AB (1995) Efficient estimation of monotone boundaries. Ann Stat 23(2):476–489

Krivonozhko VE, Utkin OB, Volodin AV, Sablin IA, Patrin M (2004) Constructions of economic functions and calculations of marginal rates in DEA using parametric optimization methods. J Oper Res Soc 55(10):1049–1058

Kuosmanen T (2001) DEA with efficiency classification preserving conditional convexity. Eur J Oper Res 132(2):326–342

Macher JT (2006) Technological development and the boundaries of the firm: a knowledge-based examination in semiconductor manufacturing. Manage Sci 52(6):826–843

McFadden DL, Fuss M (1978) Production economics: a dual approach to theory and applications. North-Holland, Amsterdam

Morita H (2003) Analysis of economies of scope by data envelopment analysis: comparison of efficient frontiers. Int Trans Oper Res 10:393–402

Panzar JC, Willig RD (1981) Economies of scope. Am Econ Rev 71(2):268–272

Paul CJM, Nehring R (2005) Product diversification, production systems, and economic performance in US agricultural production. J Econ 126(2):525–548

Preyra C, Pink G (2006) Scale and scope efficiencies through hospital consolidations. J Health Econ 25(6):1049–1068

Prior D (1996) Technical efficiency and scope economics in hospitals. Appl Econ 28(10):1295–1301

Rawley E, Simcoe TS (2010) Diversification, diseconomies of scope, and vertical contracting: evidence from the taxicab industry. Manage Sci 56(9):1534–1550

Rosko MD, Mutter RL (2011) What have we learned from the application of stochastic frontier analysis to US hospitals? Med Care Res Rev 68(1):75S–100S

Sav GT (2004) Higher education costs and scale and scope economies. Appl Econ 36(6):607–614

Smet M (2007) Measuring performance in the presence of stochastic demand for hospital services: an analysis of Belgian general care hospitals. J Prod Anal 27(1):13–29

Thiry B, Tulkens H (1992) Allowing for inefficiency in parametric estimation of production functions for urban transit frims. J Prod Anal 3:45–65

Wheelock DC, Wilson PW (2009) Robust nonparametric quantile estimation of efficiency and productivity change in US Commercial Banking, 1985–2004. J Bus Econ Stat 27(3):354–368

Wolinsky A (1986) The nature of competition and the scope of firms. J Ind Econ 34(3):247–259