European Journal of Operational Research 329 (2026) 966–980

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/eor

Decision Support

Natural divisibility bridging convex and nonconvex technologies: Bargaining-based estimation by cost and revenue functions ☆

Kristiaan Kerstens [a], Stefano Nasini [b],*, Rabia Nessah [b]

[a] Univ. Lille, CNRS, IESEG School of Management, UMR 9221 - LEM - Lille Économie Management, F-59000 Lille, France

[b] IESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM- Lille Economie Management, F-59000 Lille, France

ARTICLE INFO

JEL classification:
D24
C78

Keywords:
Natural divisibility estimation
Cost function
Revenue function
Cooperative bargaining

ABSTRACT

This contribution introduces a game-theoretic framework to infer divisibility levels from observed input–output production data. This is accomplished through a new class of M-parametrized deterministic, nonparametric technologies, which extend the conventional convex ($M = \infty$) and nonconvex ($M = 1$) alternatives by incorporating the new notion of natural divisibility. The statistical estimation of M is rooted in a bargaining game involving two hypothetical players pursuing conflicting objectives: efficiency and divisibility, where efficiency is measured in terms of cost and revenue functions (whose value is influenced by M), whereas the divisibility is measured by M itself. We employ the Kalai–Smorodinsky bargaining solution as an axiomatic approach to achieve an equilibrium divisibility level within the M-parametrized production possibility set. We conduct numerical tests using two secondary data sources, which reveal that $M = 2$ is the recurrent equilibrium divisibility. This highlights a limitation of traditional convex and nonconvex frontier methods, which both ignore the need for an endogenous assessment of natural divisibility.

1. Introduction

Production possibility sets (PPS) or technologies are commonly used to define relationships between inputs and outputs of production. Understanding and estimating input and output divisibility within PPSs is crucial for the correct assessment of production efficiency and underpinning decisions about optimal industry structure. The idea of divisibility, which determines how inputs and outputs can be scaled down, is fundamental to resolving trade-offs between efficiency and resource allocation in industrial scaling decisions (see the seminal article of Baumol and Fischer (1978) and Cesaroni (2020) for a n example of a recent study). For instance, in the energy market, the optimal scaling of wind farms or solar installations can be addressed once natural divisibility is known.

In a seminal contribution, Banker et al. (1984) and Färe et al. (1983) implicitly assume infinitely divisible inputs and outputs of production and propose to use convex combinations of activities to define the frontier of a PPS. They suggest to evaluate the performance of each observed organization with respect to the boundary of this convex PPS (henceforth CPPS). This methodology is part of what is known as the nonparametric approach to production analysis (see Varian (1984)).[1] Conversely, an alternative definition of the PPS replaces the convex combination of activities by their binary selection yielding a nonconvex PPS (introduced by Deprins et al. (1984)). Green and Cook (2004) note that this nonconvex PPS (henceforth NCPPS) is attractive because it almost always achieves the best fit to the observations, but also because from a managerial perspective, it measures the efficiency of an organization against an observed performance in the data set (rather than against some "synthetic" unobserved convex combination).[2]

Both NCPPS and CPPS postulate strong and opposite modeling assumptions with respect to the notion of divisibility of the input–output pairs. In fact, while CPPS assumes an infinite population of strictly efficient input–output pairs (all the pairs obtained as convex combinations of observed pairs), NCPPS relies on the idea that the observed organizations are an exhaustive sample of the input–output population and that all strictly efficient input–output pairs must be contained in the observed sample. Although the choice of either of these PPSs is important (perhaps even

☆ We are grateful to the three most constructive referees of this journal. The usual disclaimer applies.

* Corresponding author.
E-mail addresses: k.kerstens@ieseg.fr (K. Kerstens), s.nasini@ieseg.fr (S. Nasini), r.nessah@ieseg.fr (R. Nessah).

[1] In the operations research literature, these models are known under the name Data Envelopment Analysis (DEA).

[2] This particular PPS is sometimes known as the Free Disposal Hull (FDH).

https://doi.org/10.1016/j.ejor.2025.07.046
Received 26 July 2024; Accepted 21 July 2025
Available online 5 August 2025

crucial) in efficiency and productivity analysis, its selection remains an exogenous modeling choice in state-of-the-art approaches (see for a recent survey Sickles & Zelenyuk, 2019). In practical applications, assuming a PPS structure exogenously can lead to significant distortions in efficiency measurement. For example, in regulatory benchmarking exercises (e.g., for utilities or banks), using a CPPS may overestimate inefficiency due to the inclusion of unobserved convex combinations of actual firms. Conversely, using a NCPPS may underestimate potential improvement opportunities by restricting reference sets to observed peers only. In both cases, the efficiency assessments (and the managerial or policy recommendations that follow) can be misaligned with the true underlying technology.

This paper presents a new methodology to study discrete technological frontiers and to estimate the discretization level from observed input–output data. This is done by a family of M-parametrized PPS, containing the CPPS ($M = \infty$) and the NCPPS ($M = 1$) as extreme cases of the proposed M-parametrization. In line with the idea of Green and Cook (2004), our parametrized PPS extends the NCPPS one by including points that are obtained by a discrete weighted combination of subsets of organizations. Hence, in contrast with the NCPPS, reference points in the frontier of the PPS are not restricted to single organizations. However, in contrast with the CPPS, a discrete weighted combination rather than a convex combination is adopted. By tuning these discrete combination weights, the proposed parametrized PPS accommodates both extreme cases of CPPS and NCPPS.

The idea of a family of PPSs that nests as (restrictive) extreme cases both CPPS and NCPPS has been introduced by Chavas and Kim (2015). These authors propose a neighborhood-based approach, where subsets of activities are taken together. While sharing a common objective, a key difference between our natural divisibility based approach and the Chavas and Kim (2015) neighborhood-based approach relates to the economic interpretation of nonconvexity. Chavas and Kim (2015) assumes away convexity by enforcing neighborhood-based constraints on the aggregation of downscaled plans, capturing specialization and diseconomies of scope, which are in conflict with the idea of convex isoquants, but not with the idea of having isoquants that are unions of convex subsets of neighbor ing points. Conversely, our approach to address nonconvexity enforces constraints on the level of aggregation of downscaled plans, which reflects different input–output divisibility levels.

While Frank (1969) is likely the first economic analysis of production theory under the indivisibility of inputs and outputs, Bevia et al. (1999) is likely the first study of economic equilibrium with indivisible goods. The study of discrete divisibilities in production processes has then been refined in the microeconomic literature by the pioneering work of Scarf (1977, 1981b, 1981a, 1986, 1994). He investigates the preclusion of prices to support efficient production plans in the absence of convexity.

From the operations research perspective, the analysis of discrete divisibilities has been proposed by Lozano and Villa (2006), who address the quest for an integer-valued PPS to replace the traditional CPPS and NCPPS models by a mixed integer linear programming production frontier model that guarantees the integrality of input–output pairs. Kuosmanen and Matin (2009) emphasize that the approach of Lozano and Villa (2006) suffers from a lack of axiomatization as it violates the standard assumptions of Banker et al. (1984). Next, they propose an axiomatization for integer PPSs, compatible with free disposability. Matin and Kuosmanen (2009) extend and generalize the axiomatic foundation of Kuosmanen and Matin (2009) for the integer PPS under variable, non-decreasing and non-increasing returns to scale.

To a large extent, while the need to deal with discrete divisibilities in production processes has been integrated into the main corpus of CPPS and NCPPS by Lozano and Villa (2006), Kuosmanen and Matin (2009) and Matin and Kuosmanen (2009), the quest for endogenously defined discretizations gives rise to a whole new level of modeling complexity. This is due to the fact that discretization levels are typically unknown when production frontiers are to be estimated and are implicitly encoded by the specification of PPSs. The empirical benefit of an endogenous characterization of the discretization level can be easily appreciated in a range of productivity estimation contexts in which the underestimation and overestimation of firms' efficiency can lead to potentially opposite managerial conclusions. For instance, Bogetoft and Kerstens (2024) propose a nonparametric approach to distinguish between the usefulness and the wastefulness of the available slack in production plans, whereby the slack is associated with the input inefficiency under a specific PPS. Evidently, an endogenous characterization of the discretization levels entails a distinct PPS and in turn a distinct estimation of input inefficiency and available slacks. Related examples can be invoked for the analysis of productivity change (Atkinson et al., 2003) where the divisibility level is fixed.

Our discretization approach ranges from infinitely divisible activities in CPPS ("infinitely dense discretization") to indivisible activities in NCPPS ("binary discretization") from which an endogenous characterization can be established. By generalizing both approaches, our theoretical results show that the proposed family of parametrized PPS gives rise to a monotonic sequence of cost minimization and revenue maximization problems, leading to a monotonic sequence of efficiency estimations mapping increasing discretization levels ($M = 1$, $M = 2$, …, $M = \infty$).

Building on this fundamental property, the discretization level is endogenously inferred by setting up a system of axioms that a suitable discretization level M is required to satisfy. In this vein, we tailor a cooperative bargaining game where two imaginary players target conflicting objectives: efficiency on the one hand, and divisibility on the other hand. One player prefers a PPS in which the observed organizations attain the maximum efficiency (in casu, NCPPS); the other player prefers a PPS where all input–output pairs are infinitely divisible (in casu, CPPS). Everything in between constitutes a M-parametrized frontier that encodes a bargaining set ranging between these two extremes.

While the inefficiency estimation for nonconvex technologies has a long tradition in the mathematical economic literature (First et al., 1993), the use of bargaining models to interpret inefficiency has been initiated by Haskel and Sanchis (2000). These authors explain the observed inefficiencies by the idea that workers can bargain low effort (high crew sizes, for instance) and high wages, while firms desire exactly the opposite. The structural difference between the bargaining theory of Haskel and Sanchis (2000) and our advocated approach relies on the inferential idea of establishing an axiomatic system that a suitable divisibility level ideally should satisfy. Hence, unlike Haskel and Sanchis (2000), our bargaining approach solely pursues the inferential goal of an axiomatic characterization, in line with the work of Leamer (1978) and Ley (2006), which provide an economic micro-foundation of a Bayesian inference problem as a bargaining game over an unknown parameter value.

We adopt the Kalai–Smorodinsky Bargaining Solution (KSBS) (Kalai & Smorodinsky, 1975; Roth, 1979; Peters & Tijs, 1984) to attain an equilibrium divisibility level in the M-parametrized PPS. Next, we translate this axiomatic system into a bilevel optimization problem and solve it by Single Level Reformulation (SLR). A somewhat related approach is found in Boussemart et al. (2019) in the case of α-returns to scale, to estimate a common parameter α based on a goodness-of-fit approach. The key difference with regard to our advocated approach for the estimation of the natural divisibility relates to the monotonicity of the cost and revenue functions with respect to the unknown parameter M: this property prevents any goodness-of-fit argument from addressing the statistical estimation of M (i.e., the divisibility level closest to the data is always $M = 1$). Hence, the use of dedicated bargaining axioms allows us to circumvent the non-informativeness of the goodness-of-fit criterion in our context.

On the empirical side, the proposed methodology is tested using input–output production data from two secondary data sets. On the one hand, we employ U.S. farm sector data at the state level over the period 1960–2004 (e.g., Ball et al. (2014)). On the other hand, we exploit U.S. banking data from Aly et al. (1990) of a random sample of 322 banking firms from the year-end 1986.

Anticipating our empirical results, the U.S. farm sector data show that $M = 2$ (half-point input–output discrete combination) is the recurrent equilibrium divisibility level in the time horizon 1960–2004 when the returns to scale parameters are exogenous to the bargaining scheme (namely, when it is determined by the cost minimization or revenue maximization problems). Conversely, when the returns to scale parameters are assumed to be part of the bargaining scheme (and estimated by the KSBS approach) alongside M, the most frequent equilibrium divisibilities are $M = 2$ and $M = 3$. These numbers are stable both for cost minimization and revenue maximization approaches. An analogous result is obtained using the U.S. banking data, reaffirming the step-wise , nonconvex nature of the production possibility frontier. These findings suggest that neither the fully convex nor fully nonconvex PPS accurately reflect observed production structures. As a result, the proposed bargaining-theoretic approach to infer natural divisibility, not only gives rise to a richer theoretical framework, but also provides a pragmatic tool for practitioners who seek avoiding the risk of model misspecification, ultimately leading to more reliable benchmarking.

The rest of the contribution is organized as follows. Section 2 presents a general modeling design for cost minimization and revenue maximization based on the proposed PPS for endogenous discrete divisibility. In Section 3 we define a cooperative bargaining game between an efficiency and a divisibility player and formulate it as a bi-level optimization problem. Section 4 presents our empirical analysis. Section 5 concludes. Appendix A contains the mathematical proofs of propositions.

2. Technology, cost and revenue functions, and natural divisibility

Throughout the contribution, vectors are denoted with boldface characters and are column vectors. The symbols $\mathbb{R}$, $\mathbb{N}$, and $\mathbb{Q}$ are adopted with the usual meaning: the sets of real, natural, and rational numbers, whereas $\mathbb{R}_+$ and $\mathbb{Q}_+$ refers to the non-negative reals and rationals. The symbol $\mathbb{N}_g$ refers to all the natural numbers smaller or equal to g. All other sets are denoted with calligraphic letters.

Let $\mathcal{K}$ be a collection of firms (with $|\mathcal{K}| = K$) and for each $k \in \mathcal{K}$, let $\tilde{\mathbf{x}}_k \in \mathbb{R}^n$ and $\tilde{\mathbf{y}}_k \in \mathbb{R}^m$ be the observable input and output vectors respectively. Following a column-wise layout, we introduce the notation $Y = [(\tilde{\mathbf{y}}_1)^\top; \dots; (\tilde{\mathbf{y}}_K)^\top]^\top \in \mathbb{R}^{K\times m}$ and $X = [(\tilde{\mathbf{x}}_1)^\top; \dots; (\tilde{\mathbf{x}}_K)^\top]^\top \in \mathbb{R}^{K\times n}$, representing matrices containing the observed output and input vectors of the K firms, respectively. The universal set of observed production units is defined as $\mathcal{W} = \{(\tilde{\mathbf{x}}_1, \tilde{\mathbf{y}}_1), \dots, (\tilde{\mathbf{x}}_K, \tilde{\mathbf{y}}_K)\}$.

For a general firm, we define the PPS as follows:

$$\mathcal{T} = \{(\mathbf{x}, \mathbf{y}) \;:\; \mathbf{x} \text{ can produce at least } \mathbf{y}\}.$$

The input and output sets are $\mathcal{L}(\tilde{\mathbf{y}}) = \{\mathbf{x} \;:\; (\mathbf{x}, \tilde{\mathbf{y}}) \in \mathcal{T}\}$ and $\mathcal{P}(\tilde{\mathbf{x}}) = \{\mathbf{y} \;:\; (\tilde{\mathbf{x}}, \mathbf{y}) \in \mathcal{T}\}$, respectively. Geometrically, $\mathcal{L}(\tilde{\mathbf{y}})$ is a translation of the nonnegative orthant, whereas $\mathcal{P}(\tilde{\mathbf{x}})$ is a box in $\mathbb{R}^m$.

Traditional deterministic nonparametric frontier methods establish that PPSs must satisfy the following assumptions.

Axiom A1 (*A1: No Free Lunch*). $\mathcal{P}(0) = 0$.

Axiom A2 (*A2: Closedness*). $\mathcal{T}$ is closed.

Axiom A3 (*A3: Strong Disposability of Inputs*). If $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}$, then $(\lambda\mathbf{x}, \mathbf{y}) \in \mathcal{T}$ for all $\lambda \geq 1$.

Axiom A4 (*A4: Strong Disposability of Outputs*). If $(\mathbf{x}, \mathbf{y}) \in \mathcal{T}$, then $(\mathbf{x}, \lambda\mathbf{y}) \in \mathcal{T}$ for all $\lambda \leq 1$.

Axiom A5 (*A5: Envelopment*). $(\tilde{\mathbf{x}}_k, \tilde{\mathbf{y}}_k) \in \mathcal{T}$, for all $k \in \mathcal{K}$ (i.e., $\mathcal{W} \subseteq \mathcal{T}$).

Axioms A1 and A2 are weak mathematical regularity or technical assumptions: i.e., inaction is feasible, and technology is closed. Strong or free disposability of inputs (outputs) in Axiom A3 (Axiom A4) means that inputs (outputs) can be increased (decreased) while maintaining the same output (input) levels. By eventually weakening traditional strong disposability assumptions, one obtains technologies that allow one to model the phenomenon of congestion, which is best interpreted as an extreme form of technical inefficiency. Axiom A5 states that the technology must include all observed production units in the set $\mathcal{W}$, which is important when estimating PPSs using input–output data.

By a suitable combination of these axioms, a plethora of production models becomes available that allow estimating all traditional parameters of interest: returns to scale, economies of scale, economies of scope, etc. (see the reviews of Färe et al. (1994) and Briec et al. (2022) for the convex and nonconvex case, respectively). In this vein, a large attention has been devoted to the assumption of convexity (Banker et al., 1984). This represents the assumption of time divisibility, as it is only “valid for time divisibly-operable technologies” (Shephard (1970, p. 15)).

Axiom A6 (*A6: Convexity*). For $\lambda \in [0, 1]$, if $(\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}') \in \mathcal{T}$, then $(\lambda\mathbf{x} + (1-\lambda)\mathbf{x}', \lambda\mathbf{y} + (1-\lambda)\mathbf{y}') \in \mathcal{T}$.

Axiom A6 indicates that a linear combination of two activities is always feasible. As noted by Arrow and Hahn (1971), convexity can be established as a consequence of two more elementary hypotheses: continuous divisibility and additivity. Production is said to be continuously divisible if $\lambda\mathbf{s}$ is also an input–output pair for $\lambda \in [0, 1]$ whenever $\mathbf{s} = (\mathbf{x}, \mathbf{y})$ is an input–output pair. Production is said to be additive if $\mathbf{s}_1 + \mathbf{s}_2$ are also input–output pairs whenever $\mathbf{s}_1 = (\mathbf{x}_1, \mathbf{y}_1)$ and $\mathbf{s}_2 = (\mathbf{x}_2, \mathbf{y}_2)$ are input–output pairs.

Lemma 1 (*Lemma 1 from Arrow and Hahn (1971)*). *If production is continuously divisible and additive, then the production possibility set is convex and exhibits constant returns to scale, that is, it is a convex cone.*

Constant returns to scale is a strong assumption on technology. However, continuous divisibility is also a very strong assumption that we explicitly try to weaken by introducing a kind of natural divisibility notion. Turning to dual representations of the production possibility set, the cost and revenue functions are now defined. The cost function indicates the minimum expenditures needed to produce an output vector $\tilde{\mathbf{y}}$ for given input prices $\mathbf{p}_I \in \mathbb{R}^n$:

$$C(\tilde{\mathbf{y}}; \mathbf{p}_I) = \min\{\mathbf{p}_I^\top\mathbf{x} \;:\; (\mathbf{x}, \tilde{\mathbf{y}}) \in \mathcal{T}\} = \min\{\mathbf{p}_I^\top\mathbf{x} \;:\; \mathbf{x} \in \mathcal{L}(\tilde{\mathbf{y}})\}. \quad (1)$$

Properties of the cost function in the input prices and in the outputs are detailed in Färe (1988, p. 83, p. 87). For our purpose, it is worthwhile stressing the following property in the outputs: if the graph of the technology is convex, then the cost function $C(\tilde{\mathbf{y}};\mathbf{p}_I)$ is convex in the outputs $\tilde{\mathbf{y}}$ (Färe (1988, property C.8)). Thus, by contraposition, when the cost function is nonconvex in the outputs, then the technology is nonconvex. This implies that one ideally must test the convexity of the cost function in the outputs. The revenue function is defined as the maximum revenue obtainable from an input vector $\tilde{\mathbf{x}}$ for given output prices $\mathbf{p}_O \in \mathbb{R}^m$:

$$R(\tilde{\mathbf{x}};\mathbf{p}_O) = \max\{\mathbf{p}_O^\top \mathbf{y} \ : \ (\tilde{\mathbf{x}},\mathbf{y}) \in \mathcal{T}\} = \max\{\mathbf{p}_O^\top \mathbf{y} \ : \ \mathbf{y} \in \mathcal{P}(\tilde{\mathbf{x}})\}. \tag{2}$$

Properties of the revenue function in the output prices and in the inputs are detailed in Färe (1988, p. 93–94). A key property in the inputs for our purpose is worthwhile highlighting: if the graph of the technology is convex, then the revenue function $R(\tilde{\mathbf{x}};\mathbf{p}_O)$ is concave in the inputs $\tilde{\mathbf{x}}$ (Färe (1988, property R.8)). Again by contraposition, when the revenue function is nonconcave in the inputs, then the technology is nonconvex. This implies that one ideally must test the concavity of the revenue function in the inputs. We now turn to the specification of the traditional CPPS and NCPPS models: these are the basis to develop the new naturally divisible PPS.

Given a set of observed input–output pairs $\mathcal{W}$, let $\mathbf{z}$ be the K-dimensional activity vector attaching weights to each observation, in such a way as to impose that an efficient input $\mathbf{x}$ must be a combination of observed inputs X (for cost minimization) or an efficient output $\mathbf{y}$ must be a combination of observed output Y (for revenue maximization). In the traditional nonparametric production frontier literature, following Briec et al. (2004) the PPS induced by $\mathcal{W}$ is specified as follows:

$$\mathcal{T}_\ell(\delta) = \left\{(\mathbf{x},\mathbf{y}) \ : \ \mathbf{x} \geq \delta X^\top \mathbf{z},\ \mathbf{y} \leq \delta Y^\top \mathbf{z},\ \mathbf{z} \in \Lambda^\ell\right\}, \tag{3}$$

where δ is defined as a returns to scale factor (i.e., a proportionality constant), which is assumed to belong to a set $\Delta = \{d \in \mathbb{R}_+ \ : \ d \geq \Delta_{LB}, d \leq \Delta_{UB}\}$ with $1 \in \Delta$ and $\Delta_{LB} \leq \Delta_{UB}$. Based on the standard classification, depending on whether $\Delta_{LB} = 1$ and $\Delta_{UB} = \infty$, (i.e., $\delta \geq 1$) or $\Delta_{UB} = 1$ and $\Delta_{LB} = 0$, (i.e., $\delta \leq 1$) the PPS reflects non-decreasing or non-increasing return to scale, respectively. The scalar $\Delta_{LB} = \Delta_{UB} = 1$ denotes variable or flexible returns to scale. Finally, if $\Delta_{LB} = 0$ and $\Delta_{UB} = \infty$, then we have the case of constant return to scale. Additionally, $\Lambda_\ell = \left\{\mathbf{z} \in [0,1]^K \ : \ \mathbf{z}^\top \mathbf{1} = 1\right\}$, for $\ell = C$ (full divisibility), and $\Lambda_\ell = \left\{\mathbf{z} \in \{0,1\}^K \ : \ \mathbf{z}^\top \mathbf{1} = 1\right\}$, for $\ell = NC$ (non-divisibility). The labels C and NC refer to the convex and nonconvex PPS, respectively related to the CPPS and NCPPS approaches. Note that in the case of non-divisibility, the input–output of any firm must be proportional to one of those observed in $\mathcal{W}$, based on the binary selection (NCPPS approach). Conversely, in the case of full-divisibility, the input–output of any firm is proportional to any convex combinations of the ones in $\mathcal{W}$ (CPPS approach). On that account, the dividing line between convex and nonconvex frontier methods pivots on the notion of divisibility, which is assumed to be either infinity (for the convex case) or null (for the nonconvex case).

Building on $\mathcal{T}_\ell(\delta)$, linear programming and mixed-integer-linear-programming approaches can be used to characterize the cost and revenue functions $C(\tilde{\mathbf{y}};\mathbf{p}_I)$ and $R(\tilde{\mathbf{x}};\mathbf{p}_O)$, respectively (see Briec et al. (2004)). Using a standard notation for both problems, we define the following matrix objects:

$$A(s) = \begin{cases} Y^\top & \text{if } s = 1, \\ -X^\top & \text{if } s = 0, \end{cases} \qquad q(s) = \begin{cases} m & \text{if } s = 1, \\ n & \text{if } s = 0, \end{cases} \qquad \mathbf{p}(s) = \begin{cases} \mathbf{p}_I & \text{if } s = 1, \\ -\mathbf{p}_O & \text{if } s = 0, \end{cases}$$

and

$$D(s) = \begin{cases} I_n & \text{if } s = 1, \\ -I_m & \text{if } s = 0, \end{cases} \qquad \mathbf{b}_k(s) = \begin{cases} \tilde{\mathbf{y}}_k & \text{if } s = 1, \\ -\tilde{\mathbf{x}}_k & \text{if } s = 0, \end{cases} \qquad \text{and} \quad A(\delta,s) = \delta A(s).$$

where $s = 1$ refers to cost minimization and $s = 0$ refers to revenue maximization. The cost and revenue functions can be expressed as:

$$\overline{\alpha}_k(\delta_k,\ell,s) = \begin{cases} \min\limits_{\mathbf{z},\ \mathbf{w}} \quad \mathbf{p}(s)^\top \mathbf{w} & (4a) \\ \text{subj. to} \quad A(\delta_k,s)\mathbf{z} \geq \mathbf{b}_k(s) & (4b) \\ \qquad A(\delta_k,1-s)\mathbf{z} + D(s)\mathbf{w} \geq 0 & (4c) \\ \qquad \mathbf{z} \in \Lambda^\ell, & (4d) \end{cases}$$

when the return to scale factor δ_k is regarded as an exogenous parameter with respect to the cost and revenue function, and

$$\overline{\overline{\alpha}}_k(\ell,s) = \begin{cases} \min\limits_{\delta,\ \mathbf{z},\ \mathbf{w}} \quad \mathbf{p}(s)^\top \mathbf{w} & (5a) \\ \text{subj. to} \quad A(\delta,s)\mathbf{z} \geq \mathbf{b}_k(s) & (5b) \\ \qquad A(\delta,1-s)\mathbf{z} + D(s)\mathbf{w} \geq 0 & (5c) \\ \qquad \mathbf{z} \in \Lambda^\ell,\ \delta \in \Delta, & (5d) \end{cases}$$

when the return to scale factor is included as a decision variable. Note that problem (4a)–(4d) is less common than problem (5a)–(5d) to represent cost and revenue functions in the production literature. However, this representation is instrumental to the bargaining scheme to be introduced in Section 3.

Note that in light of the parameter signs and the direction of constraints, there is no need for variable bounds $\mathbf{w} \geq 0$ in problems (4a)–(4d) and (5a)–(5d). Altogether, given a collection of observed input–output pairs $\mathcal{W}$ (encoded in matrices $A(s)$, for $s \in \{0,1\}$), we obtain

$$\overline{\alpha}_k(\delta_k,\ell,s) = \begin{cases} \overline{C}(\delta_k,\tilde{\mathbf{y}};\mathbf{p}_I) & \text{if } s = 1, \\ \overline{R}(\delta_k,\tilde{\mathbf{x}};\mathbf{p}_O) & \text{if } s = 0, \end{cases} \quad \text{and} \quad \overline{\overline{\alpha}}_k(\ell,s) = \begin{cases} \overline{\overline{C}}(\tilde{\mathbf{y}};\mathbf{p}_I) & \text{if } s = 1, \\ \overline{\overline{R}}(\tilde{\mathbf{x}};\mathbf{p}_O) & \text{if } s = 0. \end{cases}$$

Note that the basic nonconvex cost function with variable returns to scale is equivalent to the Weak Axiom of Cost Minimization of Varian (1984) (see Kerstens and Van de Woestyne (2021, p. 85)).

2.1. Naturally divisible production possibility sets

Building on the characterization of $\mathcal{T}_\ell(\delta)$ in (3), forms of natural divisibility can be established by allowing for discrete input–output combinations of efficient points from $\mathcal{W}$. We consider the following natural divisibility axiom.

Axiom A7 (*A7: Natural Divisibility*). For all $(\mathbf{x},\mathbf{y}) \in \mathcal{T}$, there exists $M \in \mathbb{N}$, such that for all $\lambda \in \mathbb{N}$, we have

$$\left(\frac{M}{M+\lambda}\mathbf{x}, \frac{M}{M+\lambda}\mathbf{y}\right) \in \mathcal{T}.$$

Axiom A7 implies that an input–output pair is M-divisible if and only if it can be rescaled by $M/(M+\lambda)$ and still be a feasible input–output combination. This axiom is weaker than the traditional notion of continuous divisibility (see supra). This novel natural divisibility notion allows for the first time to ask questions about how inputs and outputs of production can be optimally scaled down. This is done by endogenizing M into a statistical estimation framework (as studied in Section 3).

We can now define the Naturally Divisible Production Possibility Set (NDPPS) as follows:

$$\mathcal{T}(\delta, M) = \left\{ (\mathbf{x},\mathbf{y}) \ : \ \mathbf{x} \geq \delta X^\top \mathbf{z},\ \mathbf{y} \leq \delta Y^\top \mathbf{z},\ \text{for } \mathbf{z} \in \left\{0, \frac{1}{M}, \frac{2}{M}, \dots, 1\right\}^K,\ \mathbf{z}^\top \mathbf{1} = 1 \right\}. \tag{6}$$

Comparing the NDPPS in (6) to the PPS in (3), we observe that both technologies contain exactly the same inequalities in inputs and outputs. However, the combination of activities that composes the envelope of input–output pairs (as established by $\mathbf{z}^\top\mathbf{1} = 1$) embodies Axiom A7 by restricting $\mathbf{z}$ to be part of the set $\left\{0, \frac{1}{M}, \frac{2}{M}, \dots, 1\right\}^K$. This entails that for each M, the set $\mathcal{T}(\delta, M)$ is nonconvex.

Remark 1. This definition implies: $\mathcal{T}(\delta,1) = \mathcal{T}_{NC}(\delta)$ and $\mathcal{T}(\delta, M) \subset \mathcal{T}_C(\delta)$, for all $M \in \mathbb{N}$.

Building on this property, this class of NDPPS nests as restrictive cases both CPPS and NCPPS, in line with the work of Chavas and Kim (2015). However, while the neighborhood-based approach proposed by Chavas and Kim (2015) consists in the combination of subsets of activities (downscaled plans, capturing specialization and diseconomies of scope), a key difference with respect to $\mathcal{T}(\delta, M)$ relates to the economic interpretation of nonconvexity. Our approach establishes a discrete aggregation of downscaled plans (which reflects different input–output divisibility levels) as microeconomic origin of nonconvexity. Hence, our advocated NDPPS does not consider the aggregation of subsets of activities, but focuses on the discrete way in which such aggregation is attained.

It is noteworthy that, unlike the modeling specifications of Lozano and Villa (2006), Kuosmanen and Matin (2009) and Matin and Kuosmanen (2009), the discrete aggregation in the proposed NDPPS does not assume input–output pairs to be integer, avoiding the need to depart from the strong disposability axioms and the creation of a dedicated collection of integer disposability axioms. The following proposition establishes natural properties of $\mathcal{T}(\delta, M)$.

Proposition 1 (*Minimum Extrapolation*). *For all $\delta \in \Delta$ and $M \in \mathbb{N}$, $\mathcal{T}(\delta, M)$ satisfies Axioms A1–A4. There exists $\delta \in \Delta$ such that $\mathcal{T}(\delta, M)$ satisfies Axiom A5.*

Proposition 2 (*Monotonicity*). *For all $\delta \in \Delta$ and $M \in \mathbb{N}$, $\mathcal{T}_{NC}(\delta) \equiv \mathcal{T}(\delta,1) \subseteq \mathcal{T}(\delta, M) \subset \mathcal{T}(\delta, 2M)$.*

Proposition 3. *Let us define*

$$\tilde{\mathcal{T}} = \bigcup_{M \in \mathbb{N}} \mathcal{T}(\delta, M).$$

$\tilde{\mathcal{T}}$ satisfies Axiom A7.

Propositions 1 and 2 reveal suitable properties that establish the NDPPS as a generalization of the traditional $\mathcal{T}_\ell(\delta)$. Proposition 3 entails that the whole NDPPS class satisfies the natural divisibility Axiom A7.

Example 1. We illustrate the NDPPS associated to a stylized case of input–output pairs. Let $\tilde{a} \in \mathbb{N}$, $\tilde{a} > 1$ and consider a collection of input–output data $\{\mathcal{W}^{(h)}\}_{h=1}^{\tilde{a}}$, with $n = m = 1$ and $K = 3$:

$$\mathcal{W}^{(h)} = \{\pi_1 = (\tilde{\mathbf{x}}_1, \tilde{\mathbf{y}}_1(h)),\ \pi_2(h) = (\tilde{\mathbf{x}}_2(h), \tilde{\mathbf{y}}_2(h)),\ \pi_3 = (\tilde{\mathbf{x}}_3, \tilde{\mathbf{y}}_3)\},$$

where for all $h = 1, \dots, \tilde{a}$, input–output pairs are defined as

$$\begin{cases} \tilde{\mathbf{x}}_1(h) = \tilde{\mathbf{y}}_1(h) = 1 \\ \tilde{\mathbf{x}}_3(h) = \tilde{\mathbf{y}}_3(h) = \tilde{a} \end{cases} \quad \text{and} \quad \left(\tilde{\mathbf{x}}_2(h),\ \tilde{\mathbf{y}}_2(h)\right) = \begin{cases} \left(\dfrac{1+\tilde{a}}{h} - \epsilon,\ 1\right) & \text{if h} > 1 \\ (\tilde{a} - \epsilon,\ 1) & \text{otherwise,} \end{cases} \tag{7}$$

for some $\epsilon > 0$. We define as $\mathcal{T}^{(h)}(\delta, M)$ the NDPPS constructed by the envelop of $\mathcal{W}^{(h)}$. Fig. 1 displays this family of NDPPSs for $\delta = 1$, where the blue, red and green dotted lines represent the frontier NDPPS with $M = 1$, $M = 2$ and $M = 3$, respectively.

The second input–output pair π_2 is output efficient with respect to $\mathcal{T}^{(h)}(1, M)$ iff

$$\exists\ \ell \in \{0, \dots, M\};\ \tilde{\mathbf{x}}_2(h) \in \left[\frac{\ell + \tilde{a}(M-\ell)}{M},\ \frac{\ell + 1 + \tilde{a}(M-\ell-1)}{M}\right] \implies \tilde{\mathbf{y}}_2(h) = \frac{\ell + \tilde{a}(M-\ell)}{M}.$$

We have the following facts:

- point $\pi_2(1)$ (blue dot in Fig. 1) is output efficient with respect to $\mathcal{T}^{(1)}(1,1)$, whereas it is not output efficient with respect to $\mathcal{T}^{(1)}(1, M)$, for any $M > 1$;

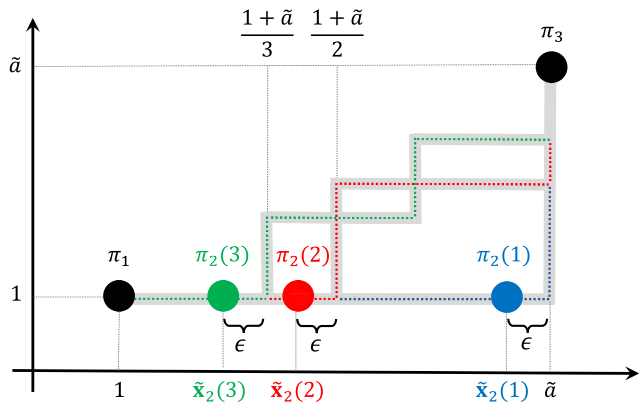


Fig. 1. Input–output pairs in $\mathcal{W}^{(1)}$, $\mathcal{W}^{(2)}$ and $\mathcal{W}^{(3)}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- point $\pi_2(2)$ (red dot in Fig. 1) is output efficient with respect to $\mathcal{T}^{(2)}(1,2)$, whereas it not output efficient with respect to $\mathcal{T}^{(2)}(1,M)$, for any $M > 2$;
- point $\pi_2(3)$ (green dot in Fig. 1), is output efficient with respect to $\mathcal{T}^{(3)}(1,3)$, whereas it is not output efficient with respect to $\mathcal{T}^{(3)}(1,M)$, for any $M > 3$.

Consequently, the efficiency or inefficiency interpretation of an input–output pair relies heavily on our assumption regarding the inherent level of natural divisibility.

2.2. Cost and revenue functions under the NDPPSs

Analogous to the problems (4a)–(4d) and (5a)–(5d) solving for the cost and revenue functions under PPS, we provide mixed-integer linear programming formulations for the cost and revenue functions under NDPPS (6) as:

$$\overline{\beta}_k(\delta_k, M, s) = \begin{cases} \min\limits_{\mathbf{z},\ \mathbf{w}} & \mathbf{p}(s)^\top \mathbf{w} & \text{(8a)} \\ \text{subj. to} & A(\delta_k, s)\mathbf{z} \geq \mathbf{b}_k(s) & \text{(8b)} \\ & A(\delta_k, 1-s)\mathbf{z} + D(s)\mathbf{w} \geq 0 & \text{(8c)} \\ & \mathbf{z}^\top \mathbf{1} = 1 & \text{(8d)} \\ & \mathbf{w} \geq 0, & \text{(8e)} \\ & z_k \in \left\{0, \frac{1}{M}, \frac{2}{M}, \dots, 1\right\}, & \text{(8f)} \end{cases}$$

when the return to scale factor δ_k is regarded as an exogenous parameter, and

$$\overline{\overline{\beta}}_k(M, s) = \begin{cases} \min\limits_{\mathbf{z},\ \mathbf{w},\ \delta} & \mathbf{p}(s)^\top \mathbf{w} & \text{(9a)} \\ \text{subj. to} & A(\delta, s)\mathbf{z} \geq \mathbf{b}_k(s) & \text{(9b)} \\ & A(\delta, 1-s)\mathbf{z} + D(s)\mathbf{w} \geq 0 & \text{(9c)} \\ & \mathbf{z}^\top \mathbf{1} = 1 & \text{(9d)} \\ & \mathbf{w} \geq 0, & \text{(9e)} \\ & z_k \in \left\{0, \frac{1}{M}, \frac{2}{M}, \dots, 1\right\} & \text{(9f)} \\ & \delta \in \Delta, & \text{(9g)} \end{cases}$$

when the return to scale factor is included as a decision variable.

Remark 2 (*Value Function Generalization*). Notice that for any $k \in \mathcal{K}$, $\delta \in \Delta$ and $M \in \mathbb{N}$ for which (4a)–(4d) is feasible, we have

$$\overline{\alpha}_k(\delta, C, s) \leq \overline{\beta}_k(\delta, M, s) \leq \overline{\beta}_k(\delta, 1, s) = \overline{\alpha}_k(\delta, NC, s).$$

In line with Remarks 1, 2 asserts a fundamental consequence of the fact that CPPS and NCPPS are included in the proposed class of NDPPSs in a hierarchical manner. This constitutes an important property in the context of the statistical estimation of natural divisibility by cooperative bargaining, as discussed next in Section 3.

The subsequent two propositions establish that problem (8a)–(8f) constitute a generalization of (4a)–(4d), for the case when $\mathcal{T}_{NC}(\delta)$ or $\mathcal{T}_C(\delta)$ are replaced by $\mathcal{T}(\delta, M)$.

Lemma 2 (*Fractional Equivalence*). *For any $M \in \mathbb{N}$, consider the following problem*

$$\tilde{\beta}_k(\delta, M, s) = \begin{cases} \min\limits_{\mathbf{z},\ \mathbf{w}} \quad \mathbf{p}(s)^\top \mathbf{w} & \text{(10a)} \\ \textit{subj. to} \quad \frac{1}{M} A(\delta, s)\mathbf{z} \geq \mathbf{b}_k(s) & \text{(10b)} \\ \qquad \frac{1}{M} A(\delta, 1-s)\mathbf{z} + D(s)\mathbf{w} \geq 0 & \text{(10c)} \\ \qquad \mathbf{z}^\top \mathbf{1} = M & \text{(10d)} \\ \qquad z_k \in \mathbb{Q}_M^K & \text{(10e)} \\ \qquad \mathbf{w} \geq 0, & \text{(10f)} \end{cases}$$

where $\mathbb{Q}_M = \{aM$ such that $a \in \mathbb{Q} \cap [0,1]\}$. There exists $M' \in \mathbb{N}$ such that for any optimal solution of (10a)–(10f) (say $\mathbf{z}(M')$), there exists an optimal solution of (8a)–(8f) (say $\mathbf{z}'(M')$), such that $\mathbf{z}(M') = M'\mathbf{z}'(M')$.

Proposition 4 (*Value Function Monotonicity*). *For any $k \in \mathcal{K}$, problem (10a)–(10f) satisfies*

$$\tilde{\beta}_k(M, s) \leq \tilde{\beta}_k(M-1, s).$$

Proposition 5 (*Value Function Convergence*). *For any $k \in \mathcal{K}$, problem (8a)–(8f) satisfies*

$$\lim_{M\to\infty} \overline{\beta}_k(\delta, M, s) = \overline{\alpha}_k(\delta, C, s).$$

and there exists $M_{UB} \in \mathbb{N}$, such that

$$\overline{\beta}_k(hM_{UB}, s) = \overline{\alpha}_k(\delta, C, s), \text{ for each } h \in \mathbb{N}.$$

Propositions 4 and 5 reaffirm the core message of Proposition 2, establishing a form of convergence of the sequence of PPS $\{\mathcal{T}(\delta, M)\}_{M=1}^{\infty}$, which approaches $\mathcal{T}_C(\delta)$ when M goes to infinity, while containing $\mathcal{T}_{NC}(\delta)$ when $M = 1$. The sequence of cost/revenue functions $\{\overline{\beta}_k(\delta, M, s)\}_{M=1}^{\infty}$ not only admits a monotonic subsequence (Proposition 4), but converges to the cost/revenue functions related to $\mathcal{T}_C(\delta)$ and $\mathcal{T}_{NC}(\delta)$ (consistent with Remark 2). It is important to note that Proposition 5 entails the search for a suitable M bounded within one and M_{UB}, as explained in detail in the next section.

3. Estimating natural divisibility by cooperative bargaining

We propose in this section a nonparametric estimation approach to infer the divisibility level $M \in \{1, \ldots, M_{UB}\}$ which is common within the same industry and the firm-specific return to scale level $\delta_k \in \Delta$, for $k \in \mathcal{K}$. This approach establishes an analogy between the statistical inference problem and a bargaining game over the parameter value, in line with the work of Leamer (1978) and Ley (2006). To this end, we define the gain functions:

$$\begin{cases} \overline{U} : \mathrm{dom}(\mathcal{W}) \times \mathbb{N} \times \mathbb{R}_+ \to [0,1] & \text{if } \delta_k \text{ is an exogenous parameter (problem (8a)–(8f))}, \\ \overline{\overline{U}} : \mathrm{dom}(\mathcal{W}) \times \mathbb{N} \to [0,1] : \mathbb{N} \to [0,1] & \text{if } \delta_k \text{ is a decision variable (problem (9a)–(9g))}, \end{cases}$$

which quantify the goodness of fit (measured as either cost or revenue efficiency) of the observed input–output pairs to the cost or revenue functions (8a)–(8f) and (9a)–(9g). A possible example of these gain functions is

$$\begin{cases} \overline{U}(\mathcal{W}, M, \boldsymbol{\delta}) & = 1 - \dfrac{1}{K} \displaystyle\sum_{k\in\mathcal{K}} \dfrac{\mathbf{p}(s)^\top \mathbf{b}_k(s) - \overline{\beta}_k(\delta_k, M, s)}{\tilde{a}_k(s)}, \\ \overline{\overline{U}}(\mathcal{W}, M) & = 1 - \dfrac{1}{K} \displaystyle\sum_{k\in\mathcal{K}} \dfrac{\mathbf{p}(s)^\top \mathbf{b}_k(s) - \overline{\overline{\beta}}_k(M, s)}{\tilde{\tilde{a}}_k(s)}, \end{cases} \tag{11}$$

where $\boldsymbol{\delta} = [\delta_1, \ \ldots, \ \delta_K]^\top$ and $|\tilde{a}_k(s)|$ and $\delta_k|\tilde{\tilde{a}}_k(s)|$ are upper bounds for $|\mathbf{p}(s)^\top \mathbf{b}_k(s) - \overline{\beta}_k(\delta, M, s)|$ and $|\mathbf{p}(s)^\top \mathbf{b}_k(s) - \overline{\overline{\beta}}_k(M, s)|$, respectively. Note that based on Proposition 2, these specifications of $\overline{U}(\mathcal{W}, M, \boldsymbol{\delta})$ and $\overline{\overline{U}}(\mathcal{W}, M)$ are non-increasing in $M \in \{m \ : \ m = 2^h, \ h \in \mathbb{N}\}$.[3]

Axiom A8 below establishes a desirable property for an estimator of M to satisfy.

Axiom A8 (*A8: Monotonicity of the Estimator*). Let $M^*(\mathcal{W})$ be the estimated natural divisibility for a data set $\mathcal{W}$. Given two data sets $\mathcal{W}$ and $\mathcal{W}'$, for all $M \in \mathbb{N}$:

$$\text{if } \max_{h\geq M} \overline{U}(\mathcal{W}, h, 1) \ \leq \ \max_{h\geq M} \overline{U}(\mathcal{W}', h, 1), \text{ then } \overline{U}(\mathcal{W}, M^*(\mathcal{W}), 1) \ \leq \ \overline{U}(\mathcal{W}', M^*(\mathcal{W}'), 1);$$
$$\text{if } \max_{h\geq M} \overline{\overline{U}}(\mathcal{W}, h) \ \leq \ \max_{h\geq M} \overline{\overline{U}}(\mathcal{W}', h), \text{ then } \overline{\overline{U}}(\mathcal{W}, M^*(\mathcal{W})) \ \leq \ \overline{\overline{U}}(\mathcal{W}', M^*(\mathcal{W}')).$$

The intuition behind this mild assumption relates to the fact that observing less efficiency in a data set (a larger mismatch between the cost or revenue function and the empirically observed cost or revenue values) must translate into less efficiency under the estimated value of M. Consistently, building on Axiom A8, the key premise behind the estimator $M^*(\mathcal{W})$ proposed in this section is that this parameter governs two

[3] Although examples of input–output data $\mathcal{W}$ for which $\overline{U}(\mathcal{W}, M, \boldsymbol{\delta})$ and $\overline{\overline{U}}(\mathcal{W}, M)$ are not monotonic for all $M \in \mathbb{N}$ can be found, in the vast majority of tested data sets, (11) turns out to be a monotonic function for all $M \in \mathbb{N}$.

conflicting objectives: (i) the efficiency of the observed organizations and (ii) the input–output divisibility. Given the inherent tension between these two objectives, we design an axiomatic characterization of $M^*(\mathcal{W})$ motivated by its ability to balance: (i) the goodness of fit maximization (which is attained under minimal divisibility $M = 1$), (ii) the traditional economic theory assumption of infinitely divisible inputs and outputs (which is attained under maximal divisibility $M = \infty$).

We approach this trade-off as a bargaining game between two imaginary players. One player (called the *efficiency player*) prefers a parametrization of the NDPPS in which the observed organizations attain the maximum efficiency (which is the case of the NCPPS associated to $M = 1$). The other player (called the *divisibility player*) prefers a parametrization of the NDPPS where all input–output pairs are infinitely divisible (which is the case of the CPPS associated to $M = M_{UB}$). The metaphor of two imaginary agents maximizing efficiency and divisibility is rooted in a statistical argument, so that the designed bargaining framework represents a conceptual tension between statistical evidence (favoring $M = 1$) and an apriori economic assumption (favoring $M = \infty$).[4] Everything in between is mapped into the utility possibility set $\mathcal{U}$ of the bargaining game, containing the utility levels associated with each value of M.

Formally, we consider the utility of the efficiency player:

$$\begin{cases} U^E \circ \{\overline{\beta}_{k,s}\}_{k\in\mathcal{K}} : \mathbb{N}\times\mathbb{R}_+ \to [0,1] & \text{if } \delta_k \text{ is an exogenous parameter (problem (8a)–(8f))},\\ U^E \circ \{\overline{\overline{\beta}}_{k,s}\}_{k\in\mathcal{K}} : \mathbb{N} \to [0,1] & \text{if } \delta_k \text{ is a decision variable (problem (9a)–(9g))}, \end{cases}$$

and the one of the divisibility player $U^D : \mathbb{N} \to [0,1]$. The function composition $(U^E \circ \{\overline{\beta}_{k,s}\}_{k\in\mathcal{K}})(\boldsymbol{\delta}, M)$ and $(U^E \circ \{\overline{\overline{\beta}}_{k,s}\}_{k\in\mathcal{K}})(M)$ are used as a short notation for $U^E(\overline{\beta}_1(\delta_1, M, s), \ldots, \overline{\beta}_K(\delta_K, M, s))$ and $U^E(\overline{\overline{\beta}}_1(M,s), \ldots, \overline{\overline{\beta}}_K(M,s))$.

Axiom A9 (*A9: Utility Properties*)**.** Players utilities are required to satisfy the following properties:

(i) If there exists $M' \in \mathbb{N}$ and $\boldsymbol{\delta}' \in \Delta^K$ such that $\mathbf{p}(s)^\top \mathbf{b}_k(s) = \overline{\beta}_k(\delta'_k, M', s)$ for all $k \in \mathcal{K}$ (all firms are efficient), then $(U^E \circ \{\overline{\beta}_{k,s}\}_{k\in\mathcal{K}})(\boldsymbol{\delta}', M') = 1$. Similarly, if there exists $M' \in \mathbb{N}$ such that $\mathbf{p}(s)^\top \mathbf{b}_k(s) > \overline{\overline{\beta}}_k(M', s)$, for all $k \in \mathcal{K}$ (all firms are efficient), then $(U^E \circ \{\overline{\overline{\beta}}_{k,s}\}_{k\in\mathcal{K}})(M') = 1$.

(ii) If there exists $k \in \mathcal{K}$, $M' \in \mathbb{N}$ and $\delta'_k \in \Delta$ for which $\mathbf{p}(s)^\top \mathbf{b}_k(s) > \overline{\beta}_k(\delta'_k, M', s)$, then $(U^E \circ \{\overline{\beta}_{k,s}\}_{k\in\mathcal{K}})(\boldsymbol{\delta}', M') < 1$. Similarly, if there exists $k \in \mathcal{K}$, $M' \in \mathbb{N}$ for which $\mathbf{p}(s)^\top \mathbf{b}_k(s) > \overline{\overline{\beta}}_k(M', s)$, then $(U^E \circ \{\overline{\overline{\beta}}_{k,s}\}_{k\in\mathcal{K}})(M') < 1$.

(iii) Both $(U^E \circ \{\overline{\beta}_{k,s}\}_{k\in\mathcal{K}})(\boldsymbol{\delta}, M)$ and $(U^E \circ \{\overline{\overline{\beta}}_{k,s}\}_{k\in\mathcal{K}})(M)$ are non-increasing in $M \in \{m \ : \ m = 2^h, \ h \in \mathbb{N}\}$ (see *Proposition* 2).

(iv) $U^D(M)$ is increasing in $M \in \mathbb{N}$ and $U^D(1) = 0$ and $U^D(M_{UB}) = 1$.

Properties (i), (ii) and (iii) imply that $(U^E \circ \{\overline{\beta}_{k,s}\}_{k\in\mathcal{K}})(\boldsymbol{\delta}, M)$ and $(U^E \circ \{\overline{\overline{\beta}}_{k,s}\}_{k\in\mathcal{K}})(M)$ correctly encode candidate gain functions $\overline{U}$ and $\overline{\overline{U}}$, which quantify the fit between the cost or revenue functions and the empirically observed cost or revenue values. Furthermore, these regularity conditions ensure that both utility functions correctly represent player's preferences for efficiency and divisibility. A possible specification of $(U^E \circ \{\overline{\beta}_{k,s}\}_{k\in\mathcal{K}})(\boldsymbol{\delta}, M)$ and $(U^E \circ \{\overline{\overline{\beta}}_{k,s}\}_{k\in\mathcal{K}})(M)$, satisfying properties (i), (ii) and (iii) in Axiom A8, is provided in (11). Likewise, a possible specification of $U^D(M)$, property (iv) in Axiom A8, is

$$U^D(M) = \frac{M_{UB}}{M_{UB} - 1}\left(1 - M^{-1}\right). \tag{12}$$

Definition 1. A cooperative bargaining game between efficiency and divisibility is a pair $(\mathcal{U}, \mathbf{d})$, where $\mathbf{d} = [d_E, d_D]^\top$ denotes the vector of disagreement points, corresponding to payoffs each player will get if the bargaining process fails, and $\mathcal{U}$ denotes the utility possibility set, which is defined as

$$\mathcal{U} = \overline{\mathcal{U}} = \left\{ (v_E, v_D) \ : \ v_E \le (U^E \circ \{\overline{\beta}_k\}_{k\in\mathcal{K}})(\boldsymbol{\delta}, M), \ v_D \le U^D(M), \ M \in \mathbb{N}_{M_{UB}}, \boldsymbol{\delta} \in \Delta^K \right\}, \tag{13}$$

if $\boldsymbol{\delta}$ is an exogenous parameter of problem (8a)–(8f), and

$$\mathcal{U} = \overline{\overline{\mathcal{U}}} = \left\{ (v_E, v_D) \ : \ v_E \le (U^E \circ \{\overline{\overline{\beta}}_k\}_{k\in\mathcal{K}})(M), \ v_D \le U^D(M), \ M \in \mathbb{N}_{M_{UB}} \right\}, \tag{14}$$

if $\boldsymbol{\delta}$ is a decision variable of problem (9a)–(9g).

The distinction between these two cases captures estimation assumptions about how the returns to scale can be inferred from the observed collection of input–output pairs $\mathcal{W}$.

Remark 3. We have the following relation: $\overline{\overline{\mathcal{U}}} \subseteq \overline{\mathcal{U}}$.

The proposed estimation method for the discrete divisibility M consists in characterizing an equilibrium point for $(\mathcal{U}, \mathbf{d})$ based on the Kalai–Smorodinsky axiomatic approach (Kalai & Smorodinsky, 1975; Roth, 1979; Peters & Tijs, 1984). Like the Nash bargaining solution, the Kalai–Smorodinsky bargaining solution (KSBS) relies upon the standard axioms of Pareto optimality, symmetry, and invariance to affine transformations. However, the KSBS replaces the independence of irrelevant alternatives assumption with a monotonicity assumption, consistently with our Axiom A8. This is based on the idea that increasing a player's utility possibility set with higher utility levels must always benefit that player.[5] In the

[4] These two principles are metaphorically depicted as utility-maximizing agents, even if they are not embodied by specific institutions. One may think about the *efficiency player* in terms of an industry association interested in promoting efficiency. Relatedly, one may think about the *divisibility player* as a regulatory body interested in promoting divisibility to guarantee competition.

[5] The independence of irrelevant alternatives axiom, that has frequently raised theoretical and experimental objections (Kalai & Smorodinsky, 1975; Luce & Raiffa, 1989), neglects the opportunity cost of unilaterally beneficial options. Different contributions have tried to relax or replace this axiom. For instance, Peters and Van Damme (1991) provide an equivalent characterization of the Nash bargaining solution that does not require the formal inclusion of this axiom, but that instead relies upon changes in the disagreement point.

context of selecting an equilibrium discrete divisibility value M, this monotonicity axiom captures the fact that enlarging the data set with the inclusion of highly efficient firms must increase the bargaining power of the efficiency player and consequently imply a reduction in the value of M (at equilibrium).

Definition 2 (*KSBS Characterization*)**.** Consider the best utility u_E and u_D each player can achieve in a feasible agreement:

$$\begin{aligned} u_E &= \sup u, \text{ subj. to } \exists\, v,\ (u,v) \in \mathcal{U}, \\ u_D &= \sup v, \text{ subj. to } \exists\, u,\ (u,v) \in \mathcal{U}. \end{aligned}$$

The KSBS is the maximum element in $\mathcal{U}$ on the line joining the disagreement point (d_E, d_D) and the best utility (u_E, u_D).

It is important to note that in its original formulation, as proposed by Kalai and Smorodinsky (1975), the characterization of the KSBS requires $\mathcal{U}$ to be a convex set (which is not the case of (13) and (14), as $M \in \mathbb{N}$). However, Conley and Wilkie (1991) show that Definition 2 is still a valid characterization of the KSBS if $\mathcal{U}$ is comprehensive.

Definition 3 (*Comprehensiveness*)**.** We say that $\mathcal{U}$ is $\mathbf{d}$-comprehensive iff for any $\boldsymbol{u} \in \mathcal{U}$, if $\mathbf{v} \preceq \boldsymbol{u}$ and $\mathbf{d} \preceq \mathbf{v}$, then $\mathbf{v} \in \mathcal{U}$, where $\preceq$ denotes the elementwise vector inequality.

The KSBS of $(\mathcal{U}, \mathbf{d})$ is then obtained by defining the line joining the disagreement point and the best utility (or disutility), which is

$$\left\{ \begin{bmatrix} d_E \\ d_D \end{bmatrix} + \begin{bmatrix} u_E - d_E \\ u_D - d_D \end{bmatrix} t \ :\ t \in \mathbb{R} \right\}.$$

Therefore, the maximum element in the intersection between this line and $\mathcal{U}$ is given by the solution of the following pair of problems:

$$\begin{cases} \max\limits_{M,\,\delta} \quad \min \left\{ \dfrac{(U^E \circ \{\overline{\beta}_k\}_{k\in\mathcal{K}})(\delta, M) - d_E}{u_E - d_E}, \dfrac{U^E(M) - d_D}{u_D - d_D} \right\} & \text{(15a)} \\ \text{s.t.} \quad M \in \{1, \dots, M_{UB}\},\ \delta \in \Delta^K, & \text{(15b)} \end{cases}$$

for the case of (13) (namely, when δ is regarded as an exogenous parameter with respect to the cost and revenue function, which entails that it is decided by the bargaining game), and

$$\begin{cases} \max\limits_{M} \quad \min \left\{ \dfrac{(U^E \circ \{\overline{\overline{\beta}}_k\}_{k\in\mathcal{K}})(M) - d_E}{u_E - d_E}, \dfrac{U^E(M) - d_D}{u_D - d_D} \right\} & \text{(16a)} \\ \text{s.t.} \quad M \in \{1, \dots, M_{UB}\}, & \text{(16b)} \end{cases}$$

for the case of (14) (namely, when δ is established in either the cost or revenue optimization problems).

The numerical characterization of the KSBS from problems (15) and (16) involves the solution of nonlinear integer bilevel problems with respect to decision variables $M \in \{1, \dots, M_{UB}\}$ and $\delta \in [\Delta_{LB}, \Delta_{UB}]^K$. Each evaluation requires the solution of K integer programming problems in the form of the cost and revenue functions (8a)–(8f) and (9a)–(9g).

Example 2. Let us consider the collection of input–output data $\{\mathcal{W}^{(h)}\}_{h=1}^{\tilde{a}}$, with $n = m = 1$ and $K = 3$, as presented in Example 1, where $\tilde{a} \in \mathbb{N}$, $\tilde{a} > 1$, $s = 0$ and $\mathbf{p}(s) = \mathbf{1}$, and input–output pairs are defined in accordance with (7). We have:

$$\mathbf{p}(0)^\top \mathbf{b}_k(0) = \begin{cases} -1 & \text{if } k = 1 \text{ and } k = 2, \\ -\tilde{a} & \text{if } k = 3. \end{cases}$$

Let $\overline{\beta}_k^{(h)}(1, M, 0)$ be the optimal value of (8a)–(8f), based on the input–output data $\mathcal{W}^{(h)}$, for each $h = 1, \dots, \tilde{a}$:

$$\overline{\beta}_k^{(h)}(1, M, 0) = \begin{cases} \min & -z_1 - z_2 - \tilde{a} z_3 \\ \text{s.t.} & z_1 + \tilde{x}_2(h) z_2 + \tilde{a} z_3 \geq \tilde{x}_2(h) \\ & z_1 + z_2 + \tilde{a} z_3 = 1 \\ & z_1, z_2, z_3 \in \left\{0, \dfrac{1}{M}, \dfrac{2}{M}, \dots, 1\right\}. \end{cases}$$

We have:

$$\begin{cases} \overline{\beta}_1^{(h)}(1, M, 0) = -1 \\ \overline{\beta}_3^{(h)}(1, M, 0) = -\tilde{a} \end{cases} \quad \text{and} \quad \overline{\beta}_2^{(h)}(1, M, 0) = \begin{cases} -1 & \text{if } h \leq M \\ -\left(1 - \dfrac{1}{M_2(h)}\right) + \dfrac{1}{M_2(h)} \tilde{a} & \text{otherwise.} \end{cases}$$

where

$$M_2(h) = \min \left\{ \ell \in \mathbb{N} \;\middle|\; \ell \left(\frac{\tilde{a} + 1}{h} - \epsilon \right) \leq \tilde{a} \right\}$$

Therefore, for all $M \leq h$:

$$\begin{cases} \mathbf{p}(0)^\top \mathbf{b}_1(0) - \overline{\beta}_1^{(h)}(1, M, 0) = 0 \\ \mathbf{p}(0)^\top \mathbf{b}_2(0) - \overline{\beta}_2^{(h)}(1, M, 0) = 0 \\ \mathbf{p}(0)^\top \mathbf{b}_3(0) - \overline{\beta}_3^{(h)}(1, M, 0) = 0 \end{cases}$$

Using (11), we obtain:

$$\tilde{U}^{(h)}(M) = (U^E \circ \{\overline{\beta}_k^h\}_{k\in\mathcal{K}})(1, M) = \begin{cases} 1 & \text{if } h \le M \\ 1 - \dfrac{\tilde{a}-1}{\tilde{a}\min\left\{\ell \in \mathbb{N} \;\middle|\; \ell\left(\dfrac{\tilde{a}+1}{h} - \epsilon\right) \le \tilde{a}\right\}} & \text{otherwise.} \end{cases}$$

Setting $d_D = d_E = 0$ and plugging the utilities into (15), we obtain:

$$\min\left\{\tilde{U}^{(h)}(M), U^D(M)\right\} = \begin{cases} U^D(M) & \text{if } h \le M \\ \min\left\{1 - \dfrac{\tilde{a}-1}{\tilde{a}\min\left\{\ell \in \mathbb{N} \;\middle|\; \ell\left(\dfrac{\tilde{a}+1}{h} - \epsilon\right) \le \tilde{a}\right\}}, U^D(M)\right\} & \text{otherwise.} \end{cases}$$

Since by Axiom A8, the utility $U^D(M)$ is non-increasing, then $\max_{\ell\in\mathbb{N}} \min\left\{\tilde{U}^{(h)}(\ell), U^D(\ell)\right\} = h$.

The main insight from Example 2 can be summarized in the following Remark 4, emphasizing that the proposed bargaining game and corresponding KSBS solution for the estimation of the natural divisibility M captures the observed discretization of input–output data $\mathcal{W}$ and is sensitive to its specific realization.

Remark 4. For each $M \in \mathbb{N}$ and for each specification of $U^D(M)$ satisfying Axiom A8, there exist an input–output data set $\mathcal{W}$ for which $M = \max_{\ell\in\mathbb{N}} \min\left\{(U^E \circ \{\overline{\beta}_k\}_{k\in\mathcal{K}})(1, \ell),\ U^D(M)\right\}$.

In the general case, since $(U^E \circ \{\overline{\beta}_k\}_{k\in\mathcal{K}})(\delta, M)$ and $(U^E \circ \{\overline{\overline{\beta}}_k\}_{k\in\mathcal{K}})(M)$ are implicitly defined from the solutions of problems (8a)–(8f) and (9a)–(9g), respectively, problems (15) and (16) constitute bilevel programming problems, where the leader selects M and δ (or M only, for the case of (16)) and K followers select $\mathbf{z}_k$ (or $\mathbf{z}_k$ and δ_k, for the case of (16)), for each $k \in \mathcal{K}$.

Note that bilevel programming techniques have been previously adopted in the context of efficiency and productivity analysis. Wu et al. (2016), for example, study operational efficiency improvement in a supply chain with an upstream leader and downstream followers to assess the potential gains from a merger. To the best of our knowledge, our own bilevel bargaining framework has not been pursued in an efficiency and productivity context.

4. Numerical and empirical analysis

This section provides empirical support to the theoretical framework developed thus far. Recall that the numerical characterization of the KSBS from problems (15) and (16) requires solving nonlinear integer bilevel problems with respect to decision variables M and δ. Furthermore, each evaluation requires the solution of K integer programming problems for the cost and revenue functions (8a)–(8f) and (9a)–(9g). One option is to construct a single-level reformulation of the nonlinear integer bilevel problems: this strategy is developed in Appendix B. An alternative option is using an enumerative method by constructing a grid enumerating over both M and δ. Depending on especially the size of the sample K, this enumerative approach may well be the preferred solution method and it is the one adopted here.

In this section, we perform a large scale computational experiment with multiple configurations of cost and revenue functions and bargaining specifications based on two empirical applications. These empirical applications are described in detail in Section 4.1 below, while the corresponding numerical tests are provided in Section 4.2.

4.1. Data

We describe here the two secondary data sets used to conduct our numerical and empirical analysis. The availability of the data sets guarantees the replicability of our results. The first data set of USA state-level agricultural panel data are publicly available at the following web address: https://github.com/StefanoNasini/Natural-Divisibility-Files/tree/main/AgriculturalData. The second data set of USA banking data is publicly available at the following web address: https://github.com/StefanoNasini/Natural-Divisibility-Files/tree/main/BankingData.

Our first application relies on USA state-level agricultural panel data that have been compiled by the U.S. Department of Agriculture (USDA). The balanced panel data set of 45 states covers the years ranging from 1960 to 2004. It includes prices and quantities for three outputs (crops, livestock, and others) and four inputs (land, intermediate inputs, capital, and labor). More details on these data are found in Ball et al. (2014). Since these agricultural data are subject to productivity change (see, e.g., Ang and Kerstens (2017) or Ball et al. (2014)), for our illustrative purpose we estimate a frontier on a year-by-year basis.

Our second application relies on a USA banking data set generated from a random sample of 322 independent banking firms from the Federal Deposit Insurance Corporation tapes for 1986 (see Aly et al. (1990)). The inputs are labor, capital, and loanable funds. Labor is measured by the number of full-time employees at the end of the time period. Capital is measured by the book value of premises and fixed assets (including capitalized leases). Loanable funds include time and saving deposits, notes and debentures, and other borrowed funds. Both capital and loanable funds are expressed in thousands of dollars. There are five outputs available measured in thousands of dollars: real estate loans; commercial and industrial loans; consumer loans; all other loans; and demand deposits. The input price of labor is obtained by dividing the total employee expenditures by the total number of employees. The input price of capital is proxied by taking the total expenditures on premises and fixed assets divided by book value. The price of loanable funds is derived by taking the sum of interest expenses on time deposits and other loanable funds divided by the total loanable funds.

Table 1
Frequency table of 45 years ×4 model instances. Rows contain the absolute frequencies of the equilibrium M over 45 instances corresponding to the time periods.

Firm problem s	δ_k selection	M			
		1	2	3	>3
rev-max $s = 0$	Model (15)	0	44	1	0
rev-max $s = 0$	Model (16)	0	9	36	0
cost-min $s = 1$	Model (15)	0	45	0	0
cost-min $s = 1$	Model (16)	0	25	20	0

Table 2
Efficiency comparison for the cost function ($s = 1$) averaged over the time horizon (1960–2004).

	$M = \infty$	$M = 2$	$M = 1$	Cost function $\bar{\bar{\beta}}_k(M, 1)$		
				min	mean	max
$M = \infty$	38.40	32.69	26.71	13 026.00	906 416.30	4 416 871.00
$M = 2$	–	32.64	25.98	22 137.37	1 074 700.50	4 416 871.00
$M = 1$	–	–	26.71	22 137.37	1 169 455.50	4 419 453.00

Table 3
Efficiency comparison for the revenue function ($s = 0$) averaged over the time horizon (1960–2004).

	$M = \infty$	$M = 2$	$M = 1$	Revenue function $\bar{\bar{\beta}}_k(M, 0)$		
				min	mean	max
$M = \infty$	28.62	9.91	1.49	35 983.77	2 509 586.00	9 826 805.00
$M = 2$	–	9.91	1.49	30 156.28	2 237 567.00	9 688 275.00
$M = 1$	–	–	1.49	30 156.28	2 116 853.00	9 642 144.00

4.2. Empirical tests

USA state-level agricultural panel data. Our first numerical tests involve a collection of 2×45 bargaining problems, obtained by solving (15) and (16) for both cost ($s = 1$) and revenue ($s = 0$) functions for each of the 46 year annual input–output data. This results in 180 numerical instances. Table 1 summarizes the M solutions of this collection of numerical instances, focusing on displaying the differences of the estimated natural divisibility under the cost minimization and revenue maximization, as well as under the endogenous bargaining selection of the returns to scale parameter.

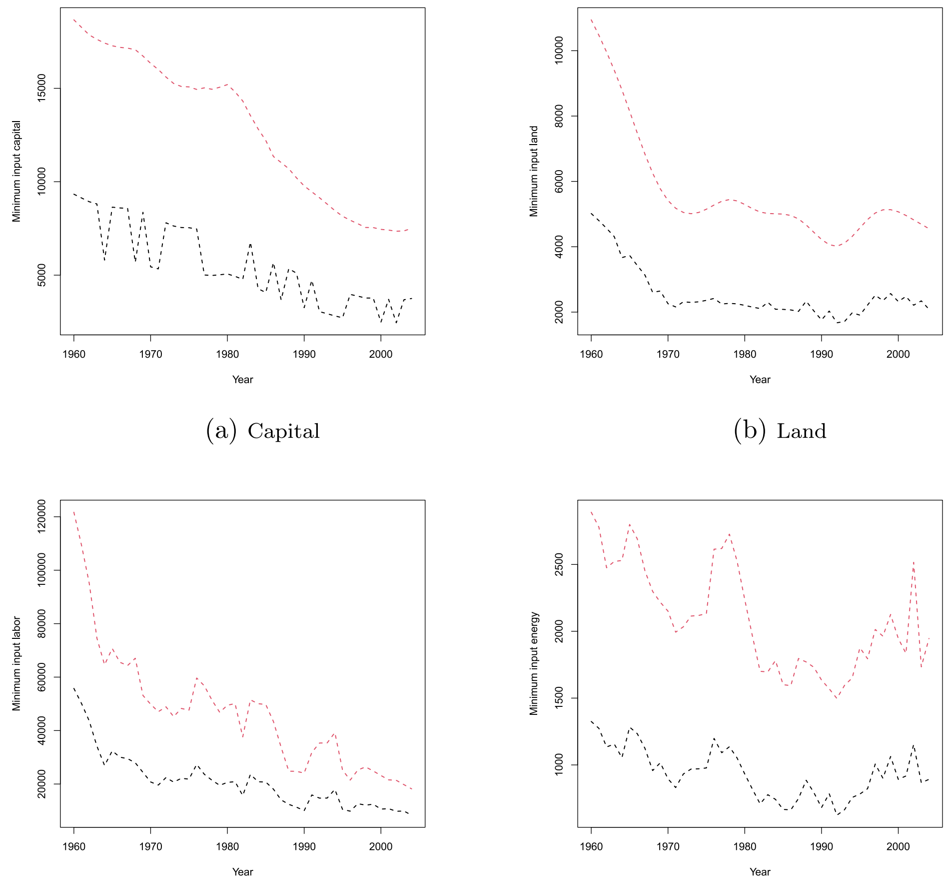
The first insight from Table 1 is the prominent appearance of $M = 2$ as the estimated natural divisibility level in the vast majority of instances. In the same line, $M = 3$ appears 57 times out of 180. No structural differences can be observed between rev-max $s = 0$ and cost-min $s = 1$. The use of model (15) leads to an overwhelming majority of $M = 2$, while the use of model (16) leads to a combination of $M = 2$ and $M = 3$.[6]

Tables 2 and 3 compare the number of inefficient states when using $M = \infty$ (CPPS), $M = 1$ (NCPPS), and when using the estimated $M = 2$ averaged over the analyzed time horizon. Columns two, three and four contain the number of states classified as inefficient by CPPS, by NDPPS, and by NCPPS. Columns five, six and seven contain the minimum, average and maximum cost functions. All values are averaged over the analyzed time horizon. Focusing first on the cost minimization in Table 2, out of 45 states, on average 38.40 are labeled as inefficient by CPPS, 26.71 by NCPPS, and 32.64 when the endogenous NDPPS is taken into account. This average tends to decline with the value of M. The number of states classified as inefficient by both CPPS and the endogenous natural divisibility approach is on average 32.69, and the number of states that are classified as inefficient by both NCPPS and the endogenous natural divisibility approach is 25.98. Finally, the number of inefficient states based on both NCPPS and CPPS is 26.71. The second vertical part of this table (columns five to seven) contains the minimum, average, and maximum cost function values for CPPS, NDPPS and NCPPS. While the average value tends to decrease in M, the maximum is identical for all M and the minimum is identical for NDPPS and NCPPS.

These numbers vary somewhat when focusing on revenue maximization in Table 3. In this case, on average 28.62 states are labeled as inefficient by CPPS, 1.49 by NCPPS, and 9.91 by the endogenous NDPPS approach. Again, this average diminishes with the value of M. Furthermore, on average 9.91 states are classified as inefficient by both CPPS and the endogenous natural divisibility approach; 1.49 states are classified as inefficient by both NCPPS and the endogenous natural divisibility approach; and 1.49 are classified as inefficient by both NCPPS and CPPS. The second vertical part of this table contains the minimum, average, and maximum revenue function values for CPPS, NDPPS and NCPPS. While the average value tends to increase in M, the maximum is identical for all M and the minimum is identical for NDPPS and NCPPS.

The summary statistics in columns five, six and seven of Table 2 report the minimum, mean, and maximum cost for $M = \infty$ (CPPS), $M = 1$ (NCPPS), and when using the estimated $M = 2$. The average cost under NCPPS is about 11% higher than under CPPS and about 3.8% higher than under $M = 2$. The minimum cost under NCPPS equals the one under $M = 2$, and both of these are 41.8% higher than under CPPS. The maximal costs under NCPPS equals the one under $M = 2$, and both of these are about 9% higher than under CPPS. Furthermore, the summary statistics of the revenue function, presented in columns five, six, and seven of Table 3, provide the minimum, mean, and maximum revenue for three scenarios: $M = \infty$ (CPPS), $M = 1$ (NCPPS), and the estimated $M = 2$. On average, revenue under CPPS is approximately 15% higher than under NCPPS and 5.73% higher than under $M = 2$. The minimum revenue in CPPS exceeds that of NCPPS by 53.76%, while NCPPS revenue is equal to that under $M = 2$. Similarly, the maximum revenue under CPPS is 4.89% higher than in NCPPS, with NCPPS revenue once again matching that under $M = 2$.

[6] To complement the assessment of this estimation results, Appendix C provides graphical illustrations of the estimated utility possibility sets, for the different values of M.



(a) Capital (b) Land

(c) Labor (d) Energy

Fig. 2. Minimum inputs for capital (a), land (b), labor (c), and energy (d). The red dashed line depicts the observed minimum input per single year, whereas the black dashed line depicts the minimum input per single year divided by the estimated divisibility level $M^*(\mathcal{W})$ of the corresponding year. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For reasons of space, we now turn to an interpretation of these optimal divisibility levels $M^*(\mathcal{W})$ for just all four inputs. Fig. 2 illustrates the extent to which the inputs capital (a), land (b), labor (c), and energy (d) can be scaled down based on the estimated divisibility level of each year along the analyzed time horizon. All four parts of the plots report: on the one hand, the observed minimum input quantity per year (in red), and on the other hand, the observed minimum input quantity divided by the optimal estimated natural divisibility level $M^*(\mathcal{W})$ (in black). Note that since we have a total of four models (revenue maximization versus cost minimization; endogenous versus exogenous δ_k), we take the average across these four optimal estimated natural divisibility levels $M^*(\mathcal{W})$.[7]

First, when analyzing the input capital in part (a) of Fig. 2, the minimal observed input experiences an almost continuously declining trend, with only a slight exception in the early 1980s. We observe a rather similar pattern for the optimal minimal input, but with only about one-half of the input quantities, with some sawtooth fluctuations over the years. Second, when analyzing the input land in part (b) of Fig. 2, the minimal observed input experiences first a strongly declining trend followed by an almost stationary trend. We observe a similar pattern for the optimal minimal input, but with only about one-third of the input quantities. Third, when analyzing the input labor in part (c) of Fig. 2, the minimal observed input experiences an almost continuously decreasing trend. We observe a slightly less attenuated pattern for the optimal minimal input, but with only about one-half of the input quantities at the start of the period and less so near the end of the period. Finally, when analyzing the input energy in part (d) of Fig. 2, then we see a slightly declining trend, but with some clear upward spikes in the 1970s as well as in recent years.

This suggests a level of input divisibility for all four inputs along the analyzed time horizon that reveals that entry barriers in agriculture are rather low and seem to have come down a bit over time. This application thereby provides a concrete managerial interpretation of the estimated $M^*(\mathcal{W})$, since it shows how optimal divisibility affects the minimum viable scale of production thereby revealing the relevance of the step-wise, nonconvex nature of the M-parametrized production frontier for firms' operations and investment decisions.

Banking data set. Our second numerical tests involve $K = 322$ independent banking firms, where $n = 5$ inputs and $m = 3$ outputs are observed, as described above. Since we only have input prices, we only consider the cost minimization approach and the two bargaining solutions, obtained by solving (15) and (16). In both cases, the estimated natural divisibility level is $M = 2$. This clearly indicates that also in the context of banking data,

[7] This result has been validated in Appendix D using a distinct characterization of the divisibility level, which consists of computing the number of times the observed input–output pairs can be obtained as M-combinations of other pairs at the level of the technology. In line with this result, we observe that for $M = 1$ and $M = 2$, many input–output pairs are obtainable as M-combinations of other pairs.

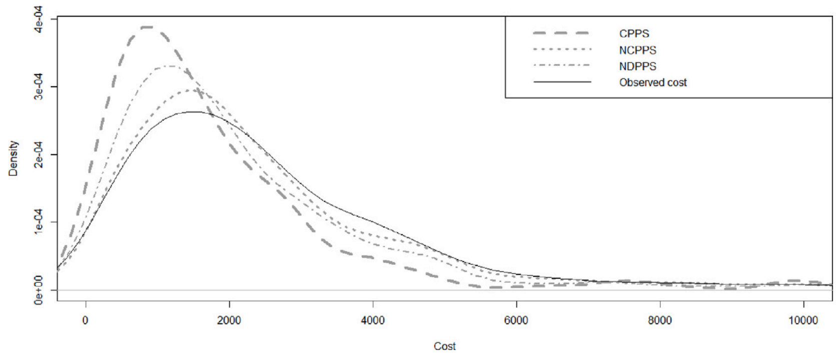


Fig. 3. Densities of banks' cost functions (truncated at 10 000) when using $M = \infty$ (CPPS), $M = 1$ (NCPPS), and when using the estimated $M = 2$ (NDPPS).

Table 4
Efficiency comparison. Columns two, three and four contain the number of banks classified as inefficient by CPPS, by NDPPS, and by NCPPS.

	$M = \infty$ (CPPS)	$M = 2$ (NDPPS)	$M = 1$ (NCPPS)	Cost function $\bar{\bar{\beta}}_k(M, 1)$		
				min	mean	max
$M = \infty$	317	263	182	24.97	2561.85	83 775.06
$M = 2$	–	264	161	41.85	2950.91	83 905.08
$M = 1$	–	–	186	49.06	3232.24	83 905.08

neither the CPPS nor the NCPPS captures the observed production structures, highlighting that real-world production technologies are neither fully continuous nor fully discretizable at the observational level, thus confirming the practical relevance of endogenizing the PPS. Table 4 provides an aggregate comparison between the number of inefficient banks when using $M = \infty$ (CPPS), $M = 1$ (NCPPS), and when using the estimated $M = 2$ (NDPPS). Correspondingly, Fig. 3 illustrates the entire distribution of banks' cost functions (truncated at 10 000), revealing substantial differences in benchmark levels across the three assumptions.

Specifically, out of 322 banking firms, 317 are labeled as inefficient by the CPPS approach, 186 by the NCPPS approach and 264 when the endogenous natural divisibility approach is taken into account. The number of firms that are classified as inefficient by both CPPS and the endogenous natural divisibility approach is 263, and the number of firms that are classified as inefficient by both NCPPS and the endogenous natural divisibility approach is 161. Finally, the number of firms classified as inefficient by both NCPPS and CPPS is 182.

The summary statistics of the cost function in columns five, six and seven of Table 4 report the minimum, mean, and maximum cost for $M = \infty$ (CPPS), $M = 1$ (NCPPS), and when using the estimated $M = 2$. The average cost under NCPPS is about 26% higher than under CPPS and about 9% higher than under $M = 2$. The minimum cost under NCPPS is 96% higher than under CPPS and about 17% higher than under $M = 2$. The maximal costs are much closer to one another.

Overall, Fig. 3 and Table 4 reaffirm that the proposed NDPPS with endogenously estimated natural divisibility situates itself in the middle point between CPPS and NCPPS, dropping the strong modeling assumptions that are opposite to one another, while being related by an extreme stance concerning the notion of divisibility of the input–output pairs. In banking regulation, this estimated divisibility level has significant benchmarking implications, as underestimating a bank's efficiency by using CPPS could result in unjustified penalties or capital requirements. Conversely, NCPPS might classify an underperforming bank as efficient, thereby distorting risk assessments.

5. Conclusions

This study presents a novel methodology for analyzing the natural divisibility of inputs and outputs and provides an estimation framework to infer divisibility levels from observed input–output data. This is critical to determine how inputs and outputs can be scaled down with direct bearing on the trade-offs between efficiency and resource allocation in industrial scaling decisions. By introducing a new class of M-parametrized deterministic nonparametric technologies, our work generalizes the traditional convex ($M = \infty$) and nonconvex ($M = 1$) alternatives, incorporating the concept of natural divisibility. Our approach mitigates the risk of underestimating potential improvement opportunities (by restricting reference sets to observed input–output pairs only, as in the NCPPS) or overestimating inefficiency (due to the inclusion of unobserved convex combinations of actual firms, as in the CPPS), thereby supporting more robust and context-sensitive efficiency evaluations.

This is done by allowing for the endogenous estimation of natural divisibility M from input–output data. This estimation relies on a cooperative bargaining game where two imaginary players target conflicting objectives: efficiency and divisibility , where efficiency is measured in terms of either cost or revenue functions (whose value is influenced by M). The KSBS is adopted as an axiomatic approach to determine an equilibrium natural divisibility level within the M-parametrized PPS.

Empirical analysis is conducted using production data from two secondary datasets: on the one hand, U.S. farm sector data at the state level from 1960 to 2004, and on the other hand U.S. banking data from a random sample of 322 bank s in 1986. The results reveal recurring equilibrium divisibility levels of $M = 2$ and $M = 3$. These findings hold for both cost minimization and revenue maximization approaches when the returns to scale parameters are considered within the bargaining scheme.

By increasing modeling flexibility, the method reveals subtler input–output dynamics, paving the way for further advancements in the field. Specifically, five further directions of research are worth pointing out. Firstly, it would be important to come up with an alternative probabilistic estimation framework to the KSBS framework which allows inferring the most likely or the expected M at the industry level.[8] Furthermore, while the current model assumes a unique M both for the inputs and outputs, a possible extension consists of considering distinct natural divisibility levels for inputs (and even for each input) and for outputs (and even for each output). In the same vein, to assess the robustness of our results, it would be crucial to determine the optimal M at the level of the PPS rather than using cost and revenue functions. Secondly, it goes without saying that to assess the validity of our M-parametrized NDPPS framework as well as its potential future extensions suggested in the previous point, it is of utmost importance that these models are replicated with alternative data sets to see whether our empirical results remain robust. To facilitate the applicability of our methodology, we provide a dedicated online repository with R codes implementing the M-divisible cost and revenue functions: https://github.com/StefanoNasini/Natural-Divisibility-Files. Thirdly, the idea of natural divisibility can also be applied to the neighborhood-based approach of Chavas and Kim (2015). This could allow focusing on two different forms of nonconvexity that can occur jointly: one driven by specialization and diseconomies of scope (captured by the selection of subsets of activities), and another driven by the discrete way in which such aggregation is attained. Fourthly, it may be worthwhile to reconsider the Bogetoft and Kerstens (2024) proposal to distinguish between useful and wasteful slack in production plans under the angle of an endogenous divisibility characterization. Finally, the inclusion of natural divisibility in the estimation of productivity change using stochastic frontiers (as introduced by Atkinson et al. (2003)) is also a further research direction.

CRediT authorship contribution statement

Kristiaan Kerstens: Writing – review & editing, Validation, Data curation, Conceptualization. **Stefano Nasini:** Writing – original draft, Software, Methodology, Formal analysis. **Rabia Nessah:** Methodology, Formal analysis.

Appendices

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ejor.2025.07.046. This material comprises four supplementary appendices: Appendix A (Mathematical proofs), Appendix B (Single-level reformulation of the Kalai–Smorodinsky bargaining problem), Appendix C (Estimated utility possibility sets), and Appendix D (Robustness validation).

References

Aly, H. Y., Grabowski, R., Pasurka, C., & Rangan, N. (1990). Technical, scale, and allocative efficiencies in US banking: An empirical investigation. *Review of Economics and Statistics*, *72*(2), 211–218.

Ang, F., & Kerstens, P. (2017). Decomposing the Luenberger-Hicks-Moorsteen total factor productivity indicator: An application to U.S. agriculture. *European Journal of Operational Research*, *260*(1), 359–375.

Arrow, K., & Hahn, F. (1971). *General competitive analysis*. Amsterdam: North-Holland.

Atkinson, S. E., Cornwell, C., & Honerkamp, O. (2003). Measuring and decomposing productivity change: stochastic distance function estimation versus data envelopment analysis. *Journal of Business & Economic Statistics*, *21*(2), 284–294.

Ball, V. E., San-Juan-Mesonada, C., & Ulloa, C. A. (2014). State productivity growth in agriculture: catching-up and the business cycle. *Journal of Productivity Analysis*, *42*(3), 327–338.

Banker, R., Charnes, A., & Cooper, W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, *30*(9), 1078–1092.

Baumol, W. J., & Fischer, D. (1978). Cost-minimizing number of firms and determination of industry structure. *Quarterly Journal of Economics*, *92*(3), 439–467.

Bevia, C., Quinzii, M., & Silva, J. A. (1999). Buying several indivisible goods. *Mathematical Social Sciences*, *37*(1), 1–23.

Bogetoft, P., & Kerstens, P. J. (2024). Distinguishing useful and wasteful slack. *Operations Research*, *72*(4), 1556–1573.

Boussemart, J.-P., Briec, W., Leleu, H., & Ravelojaona, P. (2019). On estimating optimal α-Returns to scale. *Journal of the Operational Research Society*, *70*(1), 1–11.

Briec, W., Kerstens, K., & Van de Woestyne, I. (2022). Nonconvexity in production and cost functions: An exploratory and selective review. In S. Ray, R. Chambers, & S. Kumbhakar (Eds.), *vol. 2*, *Handbook of production economics* (pp. 721–754). Singapore: Springer.

Briec, W., Kerstens, K., & Vanden Eeckaut, P. (2004). Non-convex technologies and cost functions: Definitions, duality and nonparametric tests of convexity. *Journal of Economics*, *81*(2), 155–192.

Cesaroni, G. (2020). Technically and cost-efficient centralized allocations in data envelopment analysis. *Socio-Economic Planning Sciences*, *70*, Article 100734.

Chavas, J., & Kim, K. (2015). Nonparametric analysis of technology and productivity under Non-Convexity: A neighborhood-Based approach. *Journal of Productivity Analysis*, *43*(1), 59–74.

Conley, J. P., & Wilkie, S. (1991). The bargaining problem without convexity: extending the egalitarian and Kalai-Smorodinsky solutions. *Economics Letters*, *36*(4), 365–369.

Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor efficiency in post offices. In M. Marchand, P. Pestieau, & H. Tulkens (Eds.), *The performance of public enterprises: concepts and measurements* (pp. 243–268). Amsterdam: North Holland.

Färe, R. (1988). *Fundamentals of production theory*. Berlin: Springer.

Färe, R., Grosskopf, S., & Lovell, C. (1983). The structure of technical efficiency. *Scandinavian Journal of Economics*, *85*(2), 181–190.

Färe, R., Grosskopf, S., & Lovell, C. (1994). *Production frontiers*. Cambridge: Cambridge University Press.

First, Z., Hackman, S., & Passy, U. (1993). Efficiency estimation and duality theory for nonconvex technologies. *Journal of Mathematical Economics*, *22*(3), 295–307.

Frank, C. R. (1969). *Production Theory and Indivisible Commodities*. Princeton: Princeton University Press.

Green, R. H., & Cook, W. D. (2004). A free coordination hull approach to efficiency measurement. *Journal of the Operational Research Society*, *55*(10), 1059–1063.

Haskel, J., & Sanchis, A. (2000). A bargaining model of Farrell inefficiency. *International Journal of Industrial Organization*, *18*(4), 539–556.

Kalai, E., & Smorodinsky, M. (1975). Other solutions to Nash's bargaining problem. *Econometrica*, *43*(3), 513–518.

Kerstens, K., & Van de Woestyne, I. (2021). Cost functions are nonconvex in the outputs when the technology is nonconvex: Convexification is not harmless. *Annals of Operations Research*, *305*(1–2), 81–106. http://dx.doi.org/10.1007/s10479-021-04069-1.

Kuosmanen, T., & Matin, R. K. (2009). Theory of integer-valued data envelopment analysis. *European Journal of Operational Research*, *192*(2), 658–667.

Leamer, E. E. (1978). Regression selection strategies and revealed priors. *Journal of the American Statistical Association*, *73*(363), 580–587.

Ley, E. (2006). Statistical inference as a bargaining game. *Economics Letters*, *93*(1), 142–149.

Lozano, S., & Villa, G. (2006). Data envelopment analysis of integer-valued inputs and outputs. *Computers & Operations Research*, *33*(10), 3004–3014.

Luce, R. D., & Raiffa, H. (1989). *Games and decisions: Introduction and critical survey*. Courier Corporation.

Matin, R. K., & Kuosmanen, T. (2009). Theory of integer-valued data envelopment analysis under alternative returns to scale axioms. *Omega*, *37*(5), 988–995.

Peters, H., & Tijs, S. (1984). Individually monotonic bargaining solutions for n-person bargaining games. *Methods of Operations Research*, *51*, 377–384.

Peters, H., & Van Damme, E. (1991). Characterizing the Nash and Raiffa bargaining solutions by disagreement point axioms. *Mathematics of Operations Research*, *16*(3), 447–461.

Roth, A. E. (1979). An impossibility result concerningn-person bargaining games. *International Journal of Game Theory*, *8*(3), 129–132.

Scarf, H. (1977). An observation on the structure of production sets with indivisibilities. *Proceedings of the National Academy of Sciences*, *74*(9), 3637–3641.

Scarf, H. (1981a). Production sets with indivisibilities Part I: Generalities. *Econometrica*, *49*(1), 1–32.

[8] We are grateful to an anonymous reviewer for suggesting this promising line of future research.

Scarf, H. (1981b). Production sets with indivisibilities Part II: The case of two activities. *Econometrica*, *49*(2), 395–423.
Scarf, H. (1986). Neighborhood systems for production sets with indivisibilities. *Econometrica*, *54*(3), 507–532.
Scarf, H. (1994). The allocation of resources in the presence of indivisibilities. *Journal of Economic Perspectives*, *8*(4), 111–128.
Shephard, R. (1970). *Theory of cost and production functions*. Princeton: Princeton University Press.
Sickles, R., & Zelenyuk, V. (2019). *Measurement of productivity and efficiency: Theory and practice*. Cambridge: Cambridge University Press.
Varian, H. (1984). The nonparametric approach to production analysis. *Econometrica*, *52*(3), 579–597.
Wu, D., Luo, C., Wang, H., & Birge, J. (2016). Bi-level programing merger evaluation and application to banking operations. *Production and Operations Management*, *25*(3), 498–515.