Anchor Points in DEA.

M.-L. Bougnol[1], J.H. Dulá[2]

Western Michigan University
Virginia Commonwealth University

September 2005

ABSTRACT. Anchor points play an important role in DEA theory and application. They define the transition from the efficient frontier to the "free-disposability" portion of the boundary. Our objective is to use the geometrical properties of anchor points to design and test an algorithm for their identification. We focus on the variable returns to scale production possibility set; our results do not depend on any particular DEA LP formulation, primal/dual form or orientation. Tests on real and synthetic data lead to unexpected insights into their role in the geometry of the DEA production possibility set.

Key Words: Data envelopment analysis; Linear programming; Convex analysis

---

1. Introduction

An anchor point in DEA is an extreme-efficient DMU for which some inputs can be increased and/or outputs decreased without penetrating the interior of the production possibility set. An anchor point is, therefore, an extreme element of the production possibility set that lies on the transition between the efficient frontier and the free-disposability part of the boundary.

Anchor points play an important role in DEA theory and applications. Anchor points were first named and identified by Thanassoulis and Allen (1998). They were used in the generation of unobserved DMUs created to extend the DEA efficient frontier. Anchor points also play an important role in the work of Rouse (2004) in identifying prices for healthcare services. Bougnol (2001) notes that anchor points are the only DMUs that are efficient for more than one constituency.

---

[1] ml.bougnol-potter@wmich.edu.

[2] Corresponding Author: School of Business, Virginia Commonwealth University , Richmond, VA 23284; email: `jdula@vcu.edu`.

Finally, anchor points confound DEA analysis. In the common two-stage DEA approach where the formulation forgoes non-Archimedean constants, anchor points may be confused with weak efficient DMUs.

Identifying anchor points conclusively and in general can be complicated. There exist sufficient conditions for their identification in optimal solutions of some DEA multiplier LP formulations. These conditions, may or may not be manifested, making it necessary to inspect alternate optima until either the sufficient conditions can be verified or, after exhaustive search, they never occur. Exhaustive search of alternate optima is impractical.

A specialized procedure for identifying anchor points appears in Allen and Thanassoulis (2004). The scheme applies specifically to DEA models with one input and multiple outputs under constant returns to scale and requires assumptions about the geometry of the efficient frontiers that may not always be verified. Because of this, the procedure is not guaranteed to always identify all the anchor points. We compare aspects of this procedure to ours in Appendix A.

The objective of this research is to explore the role of anchor points in the geometry of the DEA production possibility set and present and test a new algorithm for their identification. Our results suggest that anchor points play a major role in defining the shape of the DEA production possibility set.

In the next section, we introduce our notation, present a glossary of useful terms, and discuss the DMU classification scheme we will use. In Section 3, we study the geometry of anchor points and present our working definitions and the results that will make our procedure effective. In Section 4, we formalize the procedure, `AnchorProj`, for the efficient identification of anchor points. Computational results based on real and synthetic data are presented in Section 5. All proofs have been relegated to an appendix.

2. Notation and Presentation of the Model

The data consist of the set $\mathcal{A} = \{a^1, \ldots, a^n\}$ of n points, one for each DMU in the model. A DMU's data point is composed of two parts, the input components $0 \leq X_j \in \Re^{m_1}$ and the output components $0 \leq Y_j \in \Re^{m_2}$ as follows:

$$a^j = \begin{bmatrix} -X_j \\ Y_j \end{bmatrix} \in \Re^m; \quad j = 1, \ldots, n.$$

We assume that $X_j \neq 0$ and $Y_j \neq 0$ and that there is no duplication. We will focus our work on the *variable returns* ("VR") production possibility set, $\mathcal{P}^{\mathrm{VR}}$, also called the VR *hull* of the data, defined next:

$$\mathcal{P}^{\mathrm{VR}} = \left\{ z \in \Re^m \,\middle|\, z \leq \sum_j a^j \lambda_j; \text{ s.t. } \sum_j \lambda_j = 1, \ \lambda_j \geq 0; \ \forall j \right\}.$$

The set $\mathcal{P}^{\mathrm{VR}}$ is a finitely generated unbounded polyhedron that recedes at every point in the m unit directions $-e^1, \ldots, -e^{m_1}, -e^{m_1+1}, \ldots, -e^m$. The directions of recession make up a full basis for $\Re^m$ meaning the recession cone has dimension m and, therefore, $\mathcal{P}^{\mathrm{VR}}$ has full dimension.

The VR production possibility set, $\mathcal{P}^{\mathrm{VR}}$, has bounded and unbounded faces. The efficient frontier of this particular production possibility set is the union of the bounded faces. The unbounded faces make up the *free-disposability* part of the frontier. These faces recede in a positive combination of some subset of the negative unit directions. Points on the unbounded faces not in the intersection with the efficient frontier are said to be *weak efficient.* The geometry of the production possibility sets of the other three standard returns to scale assumptions (*constant returns* ("CR"), *increasing returns* ("IR"), and *decreasing returns* ("DR")) differ in important ways, e.g., all faces of the CR production possibility set are unbounded, but the general notions about anchor points presented here have equivalences. We will not, however, dwell on these cases.

We will employ a DMU classification scheme reminiscent of Charnes *et al.* (1991) based on the categories: *i)* inefficient, *ii)* efficient non-extreme, and *iii)* extreme-efficient. The three categories define three subsets of $\mathcal{A}$: 'F,' 'E,' and 'E*' respective of the order above. These three subsets are a partition of $\mathcal{A}$.

The set E* is also called the *frame* of $\mathcal{A}$. It is a minimal subset of $\mathcal{A}$ such that its VR hull is itself $\mathcal{P}^{\mathrm{VR}}$. It is easy to see that the elements of the frame correspond to extreme points of $\mathcal{P}^{\mathrm{VR}}$

(Dulá and Thrall, 2001). Frames are important because they contain all the necessary information to perform a full DEA study on the data; this can be significant when the frame is a small subset of the data points.

Before closing this section we present a short glossary of useful terms.

- The hyperplane, $\mathcal{H}(\pi, \beta) \in \Re^m$, is the set $\{y \in \Re^m | \langle \pi, y \rangle = \beta\}$. We refer to π as the *orthogonal vector* of the hyperplane and β as its *level value.*

- A *face* of a polyhedral set is the support set of a supporting hyperplane.

- A *facet* of an m-dimensional polyhedral set is an $m - 1$ dimensional face.

A DMU's classification can be conclusively resolved by solving a linear program. A DEA LP and its dual make up a "multiplier"/"envelopment" pair. The results in this paper do not depend on any particular DEA LP formulation, primal/dual form or orientation. We assume, however, that the optimal solution to the multiplier LP provides a vector, $\pi^* \in \Re^m$, and a value, $\beta^* \in \Re$, which constitute the orthogonal vector and level value of a hyperplane, $\mathcal{H}(\pi^*, \beta^*)$, that supports $\mathcal{P}^{\mathrm{VR}}$.

We explore the geometrical properties of anchor points in the next section.

3. Geometry of Anchor Points

We are interested in a particular category of elements of the frame known as *anchor points* formally defined next.

DEFINITION 1. The point $a^{\hat{j}} \in \mathrm{E}^* \subset \mathcal{A}$ is an anchor point if it belongs to an unbounded face of $\mathcal{P}^{\mathrm{VR}}$.

This definition generalizes the concept of anchor points originally introduced by Allen and Thanassoulis (2004). They worked with the specific case of one input and several outputs under constant returns to scale. Since each point can be scaled by its input value, this structure permits an analysis and visualization that is equivalent to a VR model with only outputs. As in Definition

1 above, their anchor points are a subset of E*. In their definition, an anchor point is one the outputs of which can be contracted while remaining on the free-disposability part of the boundary of the production possibility set. Our definition generalizes this concept. Anchor points, according to Definition 1, are those extreme points the inputs and outputs of which can be expanded or contracted, respectively, while remaining on the free-disposability boundary of the VR production possibility set. Refer to Appendix A for more details on the correspondence between both definitions.

Both Allen and Thanassoulis (2004) and we restrict anchor points to a subset of extreme-efficient DMUs. This restriction may be relaxed to allow all efficient DMUs to be their superset. The impact on the subsequent concepts and procedures is minor requiring a few obvious adjustments.

The concept of an anchor point presented here, along with the various results which follow, does not require that there be any specific assignment of attributes to inputs or outputs. The polyhedral set defined by nonnegative DEA data such that the inputs are negated always has the same recession cone; namely, all negative unit directions. The definition of an anchor point as an extreme point such that a translation in some negative unit direction produces a new point on a free-disposability face of the hull applies to both inputs and outputs. If the translation corresponds to an input component, the resultant point has a value for that particular input that is more negative. In terms of the actual input data, this means we are considering an effective increase in the specific input value for that point. This is the exact symmetric equivalent to a decrease in value for an output component. The relevant geometric property is that the point is at the intersection of an efficient and an unbounded face of the production possibility set.

Several interesting results and properties about anchor points are a consequence of Definition 1. We assume that the frame, E*, of the data has been extracted and all other points are momentarily discarded.

RESULT 1. *The point $a^{\hat{j}} \in$ E* is an anchor point if and only if it belongs to the support set of a supporting hyperplane $\mathcal{H}(\hat{\pi}, \hat{\beta}) \in \Re^m$ such that the orthogonal vector, $\hat{\pi}$, contains at least one zero.*

PROOF. See Appendix B.

Corollary 1 establishes a correspondence with our definition of anchor points and the one used by Thanassoulis and Allen (2004).

COROLLARY 1. *If DMU $\hat{j}$ is an anchor point then increasing an input or decreasing an output generates a new point on the free-disposability portion of the production possibility set.*

PROOF. See Appendix B.

A property of anchor points derives from their relation with simple lower dimensional projections of the production possibility set. Here is what is meant by a simple projection:

DEFINITION 2. The ith *simple projection* of $\mathcal{P}^{\mathrm{VR}}$, $\mathcal{P}_i^{\mathrm{VR}}$, is the VR production possibility set of the n data points where the ith component has been omitted.

Projected data points along with vectors and geometrical objects such as hyperplanes in the projected space will be identified by a prime "*ʹ*". The next result establishes a strong relation between points on a simple projection and anchor points.

RESULT 2. *Consider the frame of the data set,* E^*, *exclusively. The data point* $a^{\hat{j}} \in \mathrm{E}^*$ *is an anchor point if and only if it belongs to the boundary of at least one simple projection.*

Proof. See Appendix B.

Result 2 would be weaker if it applied to the entire data set and not just the frame, E^*. If the entire data set is used it is still necessary that an anchor point projects to the boundary of at least one simple projection. A sufficient condition, however, is not available since a boundary point on a projection can correspond to a weak efficient DMU in the full production possibility set. Result 2 will the basis for a procedure, `AnchorProj`, for identifying the anchor points in a DEA data set presented in the next section.

Two three-dimensional VR examples will illustrate several of the properties discussed above. In the first example, the data in Table 1 are used treating all three attributes as outputs. The frame has eight elements (extreme-efficient DMUs): $a^1, a^2, a^3, a^4, a^5, a^9, a^{10}, a^{13}$. The three simple (2-D) projections appear in Fig. 1. Fig. 1a corresponds to simple projection $i = 1$, i.e., it depicts $\mathcal{P}_1^{\mathrm{VR}}$

Table 1

Data for two DEA 3-D examples

DMUs	1	2	3	4	5	6	7	8	9	10	11	12	13
Attribute 1	4	10	13	14	5	10	3	6	3	9	12	7	9.114286
Attribute 2	12	10	6	1	9	6	7	4	3	2	1	10.5	6.757143
Attribute 3	3	4	3	5	5.5	4	6	5	8	7	2	3.5	5.578571

where only values for Attributes 2 and 3 are plotted. We can see that the projected points a^1, a^5, and a^9 are frame elements of this projected VR hull and therefore on the boundary. By Result 2, these points correspond to anchor points in the original three dimensional problem. Similarly, from Fig. 1b and 1c, we know that a^1, a^2, a^3, a^4, a^9 and a^{10} are anchor points. Table 2a summarizes these results. The letter "F" in Table 2 means that a DMU is a frame element of the VR hull and/or a projection while the letter "B" means that a DMU is a boundary point. Notice in Table 2a that DMU 13 is extreme but is not an anchor point since it does not appear on the boundary of any of the three projections.
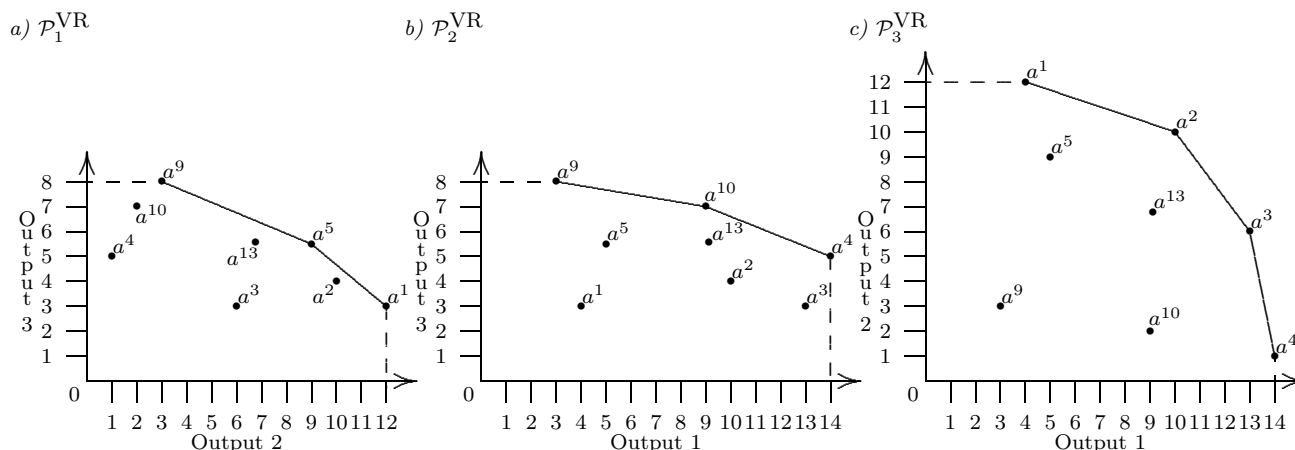

Fig. 1. Simple projections: Table 1 data treating all attributes as outputs.

A different example results from the same data in Table 1 by changing the first two attributes' roles to inputs and keeping the third as an output. The data now generate a three-dimensional VR hull with four frame elements: a^4, a^9, a^{10}, a^{11}. The three simple (2-D) projections (without negating the inputs) are displayed in Fig. 2. All four frame elements project on the boundary of the first projected VR hull, $\mathcal{P}_1^{VR}$, implying all are anchor points. Fig. 2b and 2c confirm the classification for a^9 and a^{11}. Table 2b summarizes these results. An interesting situation occurs with DMU 9 since it projects into an extreme point in all three simple projections. This is not uncommon and is verified repeatedly in our experiments with more substantial data sets. The

situation is, however, counterintuitive since in two dimensions it can only occur if there is a unique efficient DMU. Finally, it turns out that 100% of the frame elements of the full 3-D production possibility set in this example are anchor points. As we will see later, a relatively large proportion of anchor points in higher dimensions seems to be the norm.
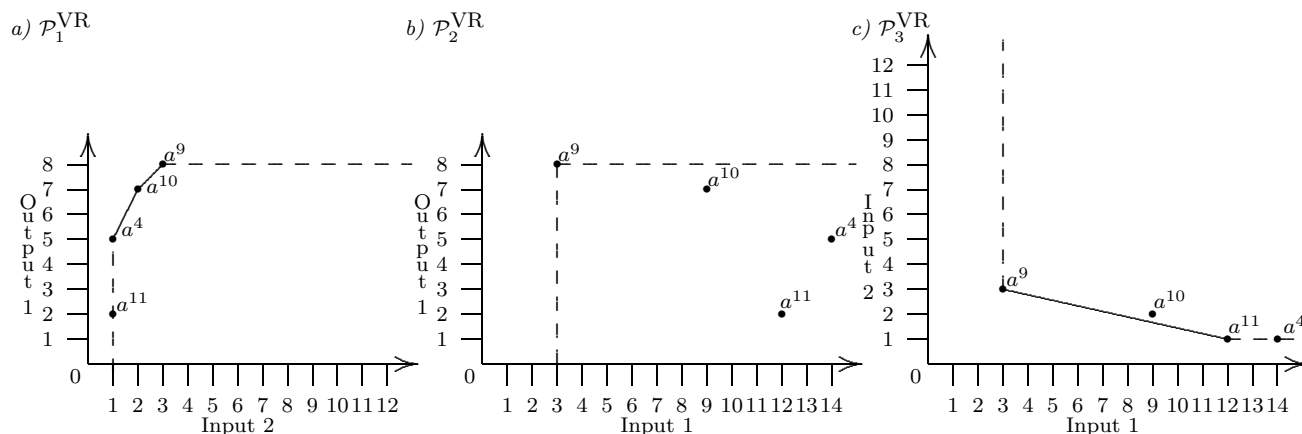

Fig. 2. Simple projections: Table 1 treating attributes 1 and 2 as inputs and attribute 3 as an output.

Table 2a

Simple projections report: 0 input, 3 outputs

DMU	Frame	Proj. 1	Proj. 2	Proj. 3
1	F	F	0	F
2	F	0	0	F
3	F	0	0	F
4	F	0	F	F
5	F	F	0	0
9	F	F	F	0
10	F	0	F	0
13	F	0	0	0

Table 2b

Simple projections report: 2 inputs, 1 output

DMU	Frame	Proj. 1	Proj. 2	Proj. 3
4	F	F	0	B
9	F	F	F	F
10	F	F	0	0
11	F	B	0	F

A note about face dimensionality and anchor points

The issue of face dimensionality arises in different aspects of DEA. It is an issue in the paper by Allen and Thanassoulis (2004). Efficient faces in DEA are bounded and they can range in dimension from $m - 1$ (facets – possibly none) to zero (extreme points – at least one). Only points in the relative interior of efficient facets have a unique supporting hyperplane that provide what Olesen and Petersen (1996) refer to as "well-defined rate of substitution". Efficient facets contain lower dimensional faces that are to be distinguished from lower dimensional faces that do not belong to any efficient facet (e.g., examples in Figures 1B, 2A, and 2B in Olesen and Petersen (1996)). We will refer to these faces as *degenerate*. Face dimensionality does not affect our notion

of anchor points or any of the results and procedures to identify them. It is interesting to note that any point on a degenerate face is on the boundary of some unbounded (i.e., free-disposability) face. To see this recall that $\mathcal{P}^{\mathrm{VR}}$ has full dimension. Therefore, any face of dimension less than $m - 1$ must be part of some higher dimensional face(s). If it is not part of a higher dimensional bounded face then it must be part of a higher dimensional unbounded face. Such a face necessarily recedes in at least one negative unit direction and is, therefore, a free disposability face. This implies that any extreme point on a degenerate face is an anchor point.

We are now ready for a formal procedure for conclusively identifying anchor points.

4. Procedures for Anchor Points

The results from the previous section provide conditions that are used in procedure `AnchorProj` to identify the anchor points of a production possibility set. This procedure is presented next.

<div style="text-align:center">

PROCEDURE `AnchorProj`

</div>

[DATA:] The set $\mathcal{A} = \{a^1, \ldots, a^n\}$.

Phase 1. Find the VR frame, E^*, of $\mathcal{A}$.

Phase 2.

Step 1. For $i = 1, \ldots, m$, find $\mathrm{E}_i^{*\prime}$; where $\mathrm{E}_i^{*\prime}$ is the subset of the frame elements that project to extreme points in simple projection i.

Step 2. Find $\mathrm{E}^{*\prime} = \cup_i^m \mathrm{E}_i^{*\prime}$.

Step 3. First Classification:

For all $a^j \in \mathrm{E}^*$ classify a^j as an anchor point if $a^j \in \mathrm{E}^{*\prime}$.

Step 4. "Mop up":

For all $a^j \in \mathrm{E}^* \setminus \mathrm{E}^{*\prime}$, test if they are on the boundary of any of the m projections. If so, classify as an anchor point; otherwise, classify as non anchor point.

END PROCEDURE `AnchorProj.`

Observations and Implementation Notes about Procedure `AnchorProj`:

1. Phase 1 requires finding the frame of the data; i.e., the extreme-efficient DMUs. This is necessary for the tests in Phase 2 to be sufficient for conclusive identification of anchor points. Recall that the frame contains all the information about $\mathcal{P}^{\mathrm{VR}}$ (Dulá and Thrall, 2001).

2. There are several ways to find the frame of a DEA data set. New efficient output sensitive algorithms for finding frame of polyhedron hulls can be found in Dulá and López (2005).

3. Frames are used in the second phase of Procedure `AnchorProj`. This is not a theoretical exigency as it is for Phase 1. In fact, the same results are obtained applying conventional DEA analysis to the m simple projections to detect boundary (efficient and weakly efficient) and interior data points in the m projected problems. The use of frames, however, has definite advantages. First, it turns out, the vast majority of extreme-efficient DMUs project into extreme points in the simple projections. Procedures specialized on finding frames extract these points efficiently. If, and when, further testing is needed to conclusively identify points that did not project into extreme points (Step 4), having the frame in hand expedites the process.

4. In our implementation, synthetic data were used to test the effect of the frame density on the performance. Non-extreme boundary points do not occur in these types of data (in general they are rare in large data sets) therefore the mop-up step (Step 4) in the procedure was not implemented with these particular data sets.

5. A "naive" approach for identifying anchor points is to perform conventional DEA analysis on the data and check for the conditions from Result 1; namely, that the DMU being scored is extreme-efficient and, if so, that at least one optimal multiplier is zero at optimality. A problem with this is that the presence of a zero multiplier in an optimal solution for an extreme-efficient DMU is sufficient for it to be an anchor point but it is not necessary. This is because extreme-efficient DMUs have multiple multiplier optimal solutions some of which may not have any zeroes even if the point is an anchor point. Looking at (possibly) all alternate optima in an LP is a notoriously difficult problem which makes a procedure relying on such searches impractical. Procedure `AnchorProj` is deterministic and conclusive. Its computational requirements are known in advance. No more than one frame identification problem for the full data set and m

additional DEA analyses for each projection need to be solved in the worse case. The approach using frames for the m subproblems proposed in Procedure `AnchorProj` does less work than that since it is opportunistic and only requires a full DEA analysis to resolve the status on non-extreme points in projected spaces. The DEA LPs used are smaller since the data are the points that correspond to the full frame in one less dimension.

6. There is a correspondence between procedure `AnchorProj` and the procedure presented by Allen and Thanassoulis (2004). This is discussed in detail in Appendix A.

Procedure `AnchorProj` was coded and tested. The results of this are reported in Section 5.

5. Computational testing of `AnchorProj`

Procedure `AnchorProj` was coded using Visual Basic for Applications and Excel. Testing was performed on an Intel Pentium M running at 1.3 GHz. The LP Solver was CPLEX8.1. Two types of tests were performed; one to assess general effectiveness of the procedure and the second specifically targeted to measure the impact of frame density; i.e., the proportion of the data points that are extreme. The first test used only real data. The second test used synthetic data generated to display different density properties. The four problems analyzed for the first test were:

1. "`Education08by70`". From Rousseau and Semple (1995).

2. "`Mortgage07by1177`". Real data from the mortgage industry; the source of this data is protected. The inputs refer to individual residence characteristics and the outputs refer to appraisal and sale values.

3. "`UFLUnivData10by628`". Real data from University of Florida *TheCenter* (Lombardi *et al.*, 2004) used in the study by Bougnol and Dulá (2005).

4. "`Banking11by19939`". Real data from the banking industry obtained from the Federal Financial Institutions Examination Council.

Table 3 reports outcomes and results for four different data sets considered to be typical of our experience. As expected, we see that times for Phase 1 grow with the cardinality of the data set.

Table 3

Experiment Results with Real Data

Data File	n	m	Percent VR Efficient	Phase 1 Time (secs.)	Phase 2 Time (secs.)	Percent Anchor Points
Education08by70	70	8 (5 inputs)	38.57	1	0	100%
Mortgage07by1177	1177	7 (5 inputs)	2.04	4	1	100%
UFLUnivData10by628	628	10 (10 outputs)	1.43	5	1	100%
Banking11by19939	19939	11 (6 inputs)	6.18	1189	447	100%

An unexpected result is the predominance of anchor points. In all real data sets, all extreme-efficient points were anchor points. The only time we witnessed a less than perfect correspondence was with the synthetic data but this did not represent an exception to this finding. In the note at the end of this section, we discuss this phenomenon in more detail.

Not so clear from this problem suite is the impact of the frame density. The frame algorithms used for these calculations are output sensitive in the sense that their performance depends on the frame density. An interesting question is how sensitive is `AnchorProj` to frame density. Fig. 3 answers this question. Recall that in Phase 1, the frame of the full data set is calculated; Phase 2 does the same work for each of the m projections. We used three different synthetic data sets, each with dimension $m = 10$ and cardinality $n = 10000$. The data sets represent three levels of frame densities: "Low" for frame densities between 1 and 5%, "Medium" for frame densities between 10 and 15% and "High" for frame densities above 20%. If we follow the time requirements for Phase 1 (black bars), we observe the expected growth in computation times that corresponds to increased frame density from "Low" to "Medium" to "High". The dramatic increase for Phase 2 computations (gray bars) is a consequence of the multiplicative effect of having to solve 10 subproblems for each data set. Anchor point preponderance (100%, 99.88%, and 99.85%, respectively) was verified with these problems too. It is clear that low frame densities can be processed quite efficiently even for large data sets.
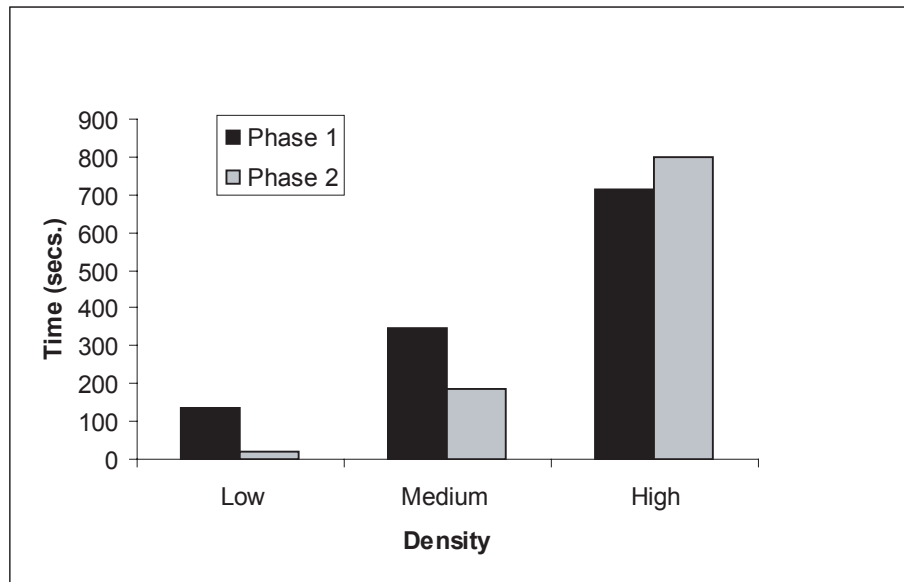
Fig. 3. Impact of Frame Density on `AnchorProj`; n=10k, m=10

A note about the prevalence of anchor points

One of the surprising results of this study is the realization that anchor points play a major role in the geometry of the production possibility set. In a large and varied problem suite which includes both real and synthetic data, it was verified, without exception, that the vast majority of all efficient DMUs are anchor points. This characteristic can also be witnessed in Allen and Thanassoulis (2004) in their application using 662 bank branches where at least 12 of the 13 efficient branches turned out to be anchors. Having used real data from various sources as well as synthetic data excludes the possibility that this phenomenon is related to the data sets.

The reasons for this anchor point prevalence are not obvious. The geometric properties of the VR polyhedral hull must play a role. A VR polyhedral hull for a DEA problem with m attributes has full dimension. There are a total of $2^m - 2 = \binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{m-1}$ receding faces ranging from edges (dimension 1) to facets (dimension $m-1$). The boundary of these faces, save the edges, can have any finite number of extreme-efficient DMUs, each of which is an anchor point. This means that $2^m - m - 2$ faces can have unlimited anchor points. Therefore, the potential for anchor points is high. For instance, in a DEA problem with ten inputs plus outputs, there are 1,022 receding faces, of which 1,012 of them can have any finite number of extreme points, i.e., anchor points. There is plenty of "room" for anchor points in a DEA VR polyhedral hull.

The paucity of non-anchor extreme-efficient DMUs illustrates how two-dimensional visualization can result in faulty intuition. In two dimensions, there are at most two anchor points and most of the 2-D examples we visualize have many more extreme-efficient DMUs. In three dimensions or more the situation suddenly inverts and anchor points become the rule in what appears to be an assault on the senses. The scarcity of non-anchor extreme-efficient DMUs suggests that the geometrical circumstances for such points to emerge are special and rare. All this serves to make the point that we are still learning and understanding the shape and nature of the DEA production possibility set.

8. Concluding remarks

Anchor points are a new category in the general classification of DMUs in DEA. They are a subset of the extreme-efficient DMUs; specifically those that are at the transition from the efficient frontier to the free-disposability part of the boundary of the production possibility set. This is the first general and comprehensive study of the role anchor points play in DEA. Their identification is a subproblem in several interesting DEA applications such as the construction of "unobserved" DMUs to capture prior value judgments in DEA, to set prices in the healthcare industry, and to identify DMUs that are efficient for multiple constituencies. The procedure, `AnchorProj` was developed for their identification based on geometric properties. This procedure was coded and applied to several large data sets. Our results show that anchor points can be efficiently identified with the use of output sensitive frame algorithms. Large data sets with 10,000 DMUs can be processed in a matter of seconds. A result from this work is the observation of an apparent preponderance of anchor points in DEA. It appears that the special properties of the VR hull make it highly likely that an efficient DMU will also be an anchor point. This realization has consequences for DEA applications that require special treatments of anchor points.

References

Allen, R., Thanassoulis, E., 2004. Improving envelopment in data envelopment analysis. European Journal of Operational Research 154 (4), 363–379.

Bougnol, Marie-Laure, 2001. Nonparametric frontier analysis with multiple constituencies. Ph.D. Dissertation, The University of Mississippi, University, MS 38677.

Bougnol, Marie-Laure, J.H. Dulá, 2004. Validating DEA as a ranking tool: An application of DEA to assess performance in higher education. To appear in Annals of Operational Research.

Charnes, A., Cooper, W.W., Thrall, R.M., 1991. A structure for classifying and characterizing efficiency and inefficiency in data envelopment analysis. The Journal of Productivity Analysis 2, 197–237.

Dulá, J.H., López, F.J., 2005. Algorithms for the frame of a finitely generated unbounded polyhedron. To appear in INFORMS Journal on Computing.

Dulá, J.H., Thrall, R.M., 2001. A computational framework for accelerating DEA. Journal of Productivity Analysis 16, 63–78.

Federal Financial Institutions Examination Council (FFIEC). In: Report of Condition and Income. `www.chicagofed.orgeconomic_research_and_datacommercial_bank_data.cfm`.

Lombardi, J., E.D. Capaldi, D.D. Craig, L.N. Collis, D.S. Gater, and D.A. Roak. (2004). "The Top American Research Universities," TheCenter's Annual Report. `www.thecenter.ufl.edu`.

Olesen, O.B., Petersen, N.C., 1996. Indicators of ill-conditioned data sets and model misspecification in data envelopment analysis: An extended facet approach. Management Science 42 (2), 205–219.

Rouse, P., 2004. Using DEA to set prices for healthcare delivery in New Zealand hospitals, Working paper. The University of Auckland, New Zealand.

Rousseau, J.J., Semple, J.H., 1995. Radii of classification preservation in data envelopment analysis: A case study of 'Program Follow-Through'. Journal of Operational Research Society 46, 943–957.

Thanassoulis, E., Allen, R., 1998. Simulating weights restrictions in data envelopment analysis by means of unobserved DMUs. Management Science 44 (4), 586–594.

Appendix A: Correspondence with the Allen-Thanassoulis Anchor Detection Procedure

Allen and Thanassoulis (2004) propose an LP formulation for detecting anchor points based on the solution of LPs for the case of the constant returns to scale with one input and multiple outputs. The procedure is adapted here for detecting anchor points in the general VR case without the use of non-Archimedean constants.

Recall that a^j is a data point composed of input $X^j \geq 0$ and output $Y^j \geq 0$ components such that $a^j = \begin{bmatrix} -X^j \\ Y^j \end{bmatrix}$. The solution to the following LP provides a sufficient condition for the identification of an anchor point, where E^* is the frame of the data set i.e., the set of extreme-efficient DMUs, and $\mathcal{S}$ is the vector of slacks and surpluses.

$$
\begin{aligned}
&\min_{\phi,\lambda\geq 0,\mathcal{S}\geq 0} \quad \phi \\
&\text{s.t.} \\
&\qquad\qquad \sum_{\{j|a^j\in\mathrm{E}^*\},j\neq j^*} a^j\lambda_j \quad - \quad I_m\mathcal{S} \quad = \quad a^{j^*} \qquad\qquad (M2') \\
&\qquad\qquad \sum_{\{j|a^j\in\mathrm{E}^*\},j\neq j^*} \lambda_j \qquad\qquad\qquad = \quad \phi
\end{aligned}
$$

LP $(M2')$ is based on LP $(M2)$ in Allen and Thanassoulis with the following differences:

1. The LP $(M2')$ is a relaxation of a non-Archimedean form.

2. The coefficient matrix involves the data for the extreme-efficient DMUs.

3. The formulation applies to the VR model.

These differences do not affect the main point of this discussion which is to show that the procedure by Allen and Thanassoulis identifies points that satisfy our definition for anchor points. We will show that any DMU j^* that generates a solution to $(M2')$ such that $\phi^* > 1$ and at least one slack is strictly positive satisfies the conditions of Result 1 for an anchor point.

By complimentary slackness, an optimal solution to $(M2')$ with at least one positive slack implies that the optimal solution, (π^*, β^*), to the dual:

$$
\begin{cases}
\max\limits_{\pi \geq 0, \beta} & \langle \pi, a^{j^*} \rangle \\
& \langle \pi, a^j \rangle \quad + \quad \beta \quad \leq \quad 0; \quad \forall j \text{ s.t. } a^j \in \mathrm{E}^*, j \neq j^* \\
& \qquad\qquad - \quad \beta \quad = \quad 1;
\end{cases}
$$

is such that at least one component of π^* is zero. Dual feasibility means that the hyperplane $\mathcal{H}(\pi^*, \beta^*)$ supports the VR hull of the extreme-efficient DMUs (without a^{j^*}) and strong duality states that this hyperplane strongly separates this hull from a^{j^*}. This is easier to see if we rewrite the dual as: (recall $\beta^* = -1$)

$$
\begin{cases}
\max\limits_{\pi \geq 0} & \langle \pi, a^{j^*} \rangle \\
& \langle \pi, a^j \rangle \quad \leq \quad 1; \quad \forall j \text{ s.t. } a^j \in \mathrm{E}^*, j \neq j^*.
\end{cases}
$$

Now the supporting and separating hyperplane is $\mathcal{H}(\pi^*, -1)$. The translated hyperplane $\mathcal{H}(\pi^*, \phi^*)$ supports the full VR hull at, and only at, a^{j^*}. By Result 1, this makes the point a^{j^*} an anchor point since the orthogonal vector contains at least one zero. ∎

The result in this appendix establishes that the points that Allen and Thanassoulis defined as anchor points also fall into this category under our definition.

Appendix B: Proofs to Theorems

RESULT 1. *The point $a^{\hat{j}} \in E^*$ is an anchor point if and only if it belongs to the support set of a supporting hyperplane $\mathcal{H}(\hat{\pi}, \hat{\beta}) \in \Re^m$ such that the orthogonal vector, $\hat{\pi}$, contains at least one zero.*

PROOF. $a^{\hat{j}}$ an anchor point means it is an extreme point of $\mathcal{P}^{\text{VR}}$ and on an unbounded face. The face is the support set of some supporting hyperplane $\mathcal{H}(\hat{\pi}, \hat{\beta})$. Such a face recedes in at least one negative unit direction. Suppose w.l.o.g., a direction of recession of this face is $-e^m$. Therefore, all points, $a(\alpha) \in \Re^m$ parameterized by the scalar $0 \le \alpha \in \Re$ as follows

$$a(\alpha) = a^{\hat{j}} - e^m \alpha$$

belong to the support set and hence to the supporting hyperplane; that is, for any $\alpha > 0$:

$$\langle \hat{\pi}, a^{\hat{j}} - e^m \alpha \rangle = -\hat{\beta};$$

$$\underbrace{\langle \hat{\pi}, a^{\hat{j}} \rangle}_{-\hat{\beta}} - \alpha \langle \hat{\pi}, e^m \rangle = -\hat{\beta};$$

$$-\alpha \langle \hat{\pi}, e^m \rangle = 0;$$

and since $e^m = (0, \cdots, 0, 1)$

$$\hat{\pi}_m = 0.$$

Since an extreme point of a polyhedral set is an extreme point of any of its facets (and vice versa) this concludes the first part of the proof.

The converse backtracks the arguments above. Consider an extreme point of a support set of a supporting hyperplane $\mathcal{H}(\hat{\pi}, \hat{\beta}) \in \Re^m$ such that the vector $\hat{\pi}$ has at least one zero component. Suppose w.l.o.g. that it is $\hat{\pi}_m = 0$. Immediately, the point must necessarily be an extreme point of the entire polyhedral set $\mathcal{P}^{\text{VR}}$. Construct the parameterized half line:

$$a(\alpha) = a^{\hat{j}} - e^m \alpha, \ \alpha \ge 0.$$

It follows from $a^{\hat{j}} \in \mathcal{P}^{\text{VR}}$ that $a(\alpha) \in \mathcal{P}^{\text{VR}}$ (any feasible point extended in a recession direction remains feasible). To see that the half line is on an unbounded face, let us explore its relation to

the supporting hyperplane $\mathcal{H}(\hat{\pi}, \hat{\beta})$:

$$\langle \hat{\pi}, a^{\hat{j}} - e^m \alpha \rangle = \underbrace{\langle \hat{\pi}, a^{\hat{j}} \rangle}_{-\hat{\beta}} - \alpha \underbrace{\langle \hat{\pi}, e^m \rangle}_{=0}$$

$$= -\hat{\beta}.$$

This places the entire half line $a(\alpha)$ in the support set of $\mathcal{H}(\hat{\pi}, \hat{\beta})$, assuring that the face is un-bounded. ∎

COROLLARY 1. *If DMU $\hat{j}$ is an anchor point then it is possible to increase an input or decrease an output and generate a new point on the free-disposability portion of the production possibility set.*

PROOF. A consequence of the existence of a halfline defined by a negative unit vector that remains on an unbounded facet of $\mathcal{P}^{\mathrm{VR}}$.

RESULT 2. *Consider the frame of the data set, E^*. The data point $a^{\hat{j}} \in \mathrm{E}^*$ is an anchor point if and only if it belongs to the boundary of at least one simple projection.*

PROOF. By Result 1, if the data point $a^{\hat{j}}$ is an anchor point then there exists a supporting hyperplane, $\mathcal{H}(\hat{\pi}, \hat{\beta})$, with orthogonal vector, $\hat{\pi}$ such that at least one component is zero. Suppose, w.l.o.g., it is $\hat{\pi}_m = 0$. The truncated vector $\hat{\pi}' = (\hat{\pi}_1, \ldots, \hat{\pi}_{m-1})$ defines a hyperplane, $\mathcal{H}(\hat{\pi}', \hat{\beta})$, in $\Re^{m-1}$. Notice that if $a^{j\,\prime}$ is the projected vector a^j without its last component then $\langle a^{j\,\prime}, \hat{\pi}' \rangle \le \hat{\beta}, \forall j$ and $\langle a^{\hat{j}}, \hat{\pi} \rangle = \hat{\beta}$. Therefore the projection of $a^{\hat{j}}$ on $\mathcal{P}_m^{\mathrm{VR}}$ is on its boundary. The converse is immediate given we are working only with frame elements. The presence of a projected extreme point, $a^{j\,\prime}$, on the boundary of, say, simple projection $\mathcal{P}_m^{\mathrm{VR}}$, means it belongs to the support set of a hyperplane with orthogonal vector $\hat{\pi}'$. A new orthogonal vector with one more dimension constructed using the same components and an extra zero for the mth component defines a supporting hyperplane in $\Re^m$ at a^j. $\hat{\pi}_m = 0$ makes a^j an anchor point. ∎