IMA Journal of Management Mathematics (2007) 1−16
doi: 10.1093/imaman/dri017

An Efficient Method for Calculating the Minimum Distance from an Operating Point to a Specific (Hyberbolic) Efficient Frontier

EDWARD A. SILVER, DIANE P. BISCHAK, GIOVANI J. C. DA SILVEIRA
Operations Management
Haskayne School of Business
University of Calgary
2500 University Drive N.W.
Calgary, Alberta T2N 1N4
Canada

This paper is concerned with movement from a current operating point so as to reach a two-dimensional, efficient frontier. After a discussion of different criteria for deciding on which point on the frontier to target, we focus, as an illustration, on a particular inventory management context and use of the criterion of minimum distance from the current point to the frontier. Specifically, the efficient frontier turns out to be an hyperbola in a two-dimensional representation of total (across a population of items) average stock (in monetary units) versus total fixed costs of replenishments per year. Any current (or proposed) operating strategy, differing from the class along the frontier, is located above the frontier. Finding the minimum distance from the current point to the frontier requires determining the smallest root of a quartic equation within a restricted range.

Keywords: efficient frontiers; exchange curves; economic order quantity; minimum distance; hyperbola

1. Introduction

Exchange curves and efficient frontiers play an important role in judging and improving the performance of organizations. In this paper we focus on a two-dimensional tradeoff, where an exchange curve shows how, under a specific operating policy, the value of one aggregate dimension of interest is downgraded as the value of the other dimension of interest is improved. Efficient frontiers represent the best that one can do in terms of an exchange curve, i.e., the value on one dimension cannot be improved without the other deteriorating. Under specific assumptions one can sometimes analytically determine the efficient frontier. Under more general circumstances the efficient frontier, or at least a tradeoff curve, can be estimated from data on the two-dimensional performance of a sample of organizational units. This can be done in terms of an input-output model, but also with respect to strategic tradeoffs such as cost-flexibility or quality-response time. A common procedure for estimating a frontier is Data Envelopement Analysis (DEA) (see, for example, Alexander et al (2003) or Chen and Zhu (2003)). More generally, statistical regression methods can be used to estimate tradeoff curves (see, for example, Noble (1995) or Domberger et al (2000)).

From a strategic standpoint management should be interested in how and where to move from the current operating position of their organization to a point on the performance frontier. In general, there is likely to be a relatively small number of points on the frontier attainable from the current operationg position of organization, each associated with a particular change of strategy and/or course of action having related one-time cost/benefits (e.g., cash flows for acquisition and installation of a new software system or for selling off excess inventory) as well as on-going cost/benefits (such as associated changes in staffing and training requirements). Ideally, one would wish to select the option that maximizes, for example, the present value of expected revenues minus costs. In practice this type of analysis is likely to be very difficult to carry out (due to stochastic, non-stationary elements and uncertainties in estimates of parameters). In this paper we focus on a simpler, surrogate measure, namely minimum distance from the current operating point to the efficient frontier. Incidentally, DEA does not directly provide such information. Schmenner and Swink (1998, p.11), in discussing a theory of performance frontiers, say "The theory ... suggests the need to develop metrics that characterize proximity to an asset frontier." Finding the closest point on the frontier should be appealing to management. Moving a minimum distance to the efficient frontier, in a sense, represents minimum departure or disturbance from current operating conditions.

In this paper we deal with the particular case of a commonly-modelled situation in inventory management where one is able to analytically determine an exchange curve which represents the efficient frontier for total average inventory (in monetary units) versus the aggregate replenishment workload (total fixed costs of replenishments per year) across a population of inventoried items. As will be seen, the current and proposed operating policies of an organization under study can be represented by points that lie above the efficient (in a minimizing sense) frontier. We specifically address how to find the point on the exchange curve that is closest to any given operating point, as well as the associated minimum distance.

In Section 2 we present the specific (economic order quantity) exchange curve which turns out to be an hyperbola. We elaborate on options for moving from an operating point to the exchange curve, including the aforementioned, minimum distance, option. Section 3 outlines the derivation of the mathematical condition implied by choosing the point on the curve that minimizes the distance from the (external) operating point. The result is a quartic (i.e., 4th order) equation whose root(s) must be obtained. Fortunately, normalization reduces the number of parameters involved, and graphical representations indicate that only one particular root is of interest. Section 4 summarizes the procedure for finding the closest point on the exchange curve and its associated distance from the current operating point. Some illustrative results are presented in Section 5 and brief conclusions are provided in Section 6. Details of mathematical derivations and formulae are included in the appendix.

2. The specific efficient frontier

Consider a population of n inventoried items numbered $i = 1, 2, \ldots, n$. We introduce the following notation:

D_i - the known, constant, demand rate (units/yr) of item i
v_i - the unit variable cost ($ or other monetary unit/unit) of item i
A_i - the fixed cost associated with each replenishment of item i
Q_i - the replenishment quantity (units) of item i
r - a common inventory carrying charge ($/$/yr)

The total relevant costs of using Q_i for item i are

$$TRC(Q_i) = \frac{Q_i v_i r}{2} + \frac{A_i D_i}{Q_i} \tag{1}$$

The economic order policy selects EOQ_i that minimizes (1), specifically

$$EOQ_i = \sqrt{\frac{2A_iD_i}{v_ir}} = \sqrt{\frac{2}{r}}\sqrt{\frac{A_iD_i}{v_i}} \tag{2}$$

The average cycle stock (expressed in \$) of item i is

$$ACS_i = \frac{EOQ_iv_i}{2} = \sqrt{\frac{1}{2r}}\sqrt{A_iD_iv_i} \tag{3}$$

and the fixed costs (again in \$) of replenishments of item i per year are

$$FCR_i = \frac{A_iD_i}{EOQ_i} = \sqrt{\frac{r}{2}}\sqrt{A_iD_iv_i} \tag{4}$$

Summing each of (3) and (4) over all items gives total average cycle stock (in \$),

$$TACS = \sqrt{\frac{1}{2r}}\sum_{i=1}^{n}\sqrt{A_iD_iv_i} \tag{5}$$

and total fixed costs of replenishments per year,

$$TFCR = \sqrt{\frac{r}{2}}\sum_{i=1}^{n}\sqrt{A_iD_iv_i} \tag{6}$$

Equations (5) and (6) are parametric equations that trace out an exchange curve of TACS versus TFCR as one changes the value of the parameter r. Multiplying (5) and (6) gives

$$(TFCR)(TACS) = \frac{1}{2}\left(\sum_{i=1}^{n}\sqrt{A_iD_iv_i}\right)^2 \tag{7}$$

The right side is just a constant for a given population of items and thus (7) represents an hyperbola when TACS is plotted against TFCR. Moreover, for any point on the curve there is an associated (implied) value of r. As shown by Silver et al (1998) this exchange curve (of TACS versus TFCR) represents an efficient frontier under the assumptions underlying the EOQ derivation, i.e., one can never operate below the curve and any policy that deviates from consistent use of economic order quantities will result in a point above the curve. Other, earlier references on EOQ exchange curves and implicit values of parameters, such as r, include Starr and Miller (1962), Eaton (1964), Brown (1967), Carlson and Therp (1977), and Gardner and Dannenbring (1979). Numerical illustrations will be provided in Section 5, but the situation is shown generically in Figure 1.

Unlike in the general context described in the Introduction, to move to any point on the EOQ exchange curve involves the same basic strategy, namely consistently using the EOQ formula of equation (2) with the same value of r for all items. It is the (implicit) choice of the value of r that defines where you will be on the curve. Thus there are an unlimited number of possible points to which to move on the curve, all using the same basic strategy. So how does one choose to which point on the curve to move? By moving to any point on the curve between P and Q inclusive, one of TFCR and TACS is improved (lowered) while the other is kept at the same value (this occurs right at P and Q) or is also improved (all points between P and Q). From a senior management or business perspective, achieving benefits on both aggregate dimensions is very attractive. This narrows the search somewhat, but how should the selection be made in this narrower range? P and Q are appealing in that the maximum reduction is achieved in one dimension while keeping the other at its value at the current, external operating point. Another possibility is to move to the closest point (S) on the curve, which will be somewhere between P and Q. As mentioned above, minimum distance can be viewed as minimum disruption of current strategy/procedures. Also, in a case where there are two or more operating points off the curve (e.g., a current operating point and one associated with another operating policy suggested by someone in the organization), the associated minimum distances are plausible measures to use to judge the relative attractiveness of the two or more (non-optimal) operating policies.

Incidentally, the costs in equation (1) are quite insensitive to deviations of Q_i away from EOQ_i (see for example, Hadley and Whitin (1963), Wagner (1975), Peterson and Silver (1979), Erkenkotter (1989)). Thus operating somewhat away from the exchange curve or somewhat inaccurately estimating the position of the curve (i.e., the constant in the right side of (7)) is not a serious issue. Hence high precision is not needed in determining the minimum distance from the current position to this particular frontier.

The next issue is how to find the point (S) on the hyperbola that is at a minimum distance from a given exterior point. We deal with this subject in the next two sections.

3. Normalization and the associated necessary condition for a minimum distance

To simplify the notation we first replace TFCR and TACS by x and y and let the coordinates of the exterior point be x_o and y_o. Thus from (7) we have

$$xy = C \tag{8}$$

where

$$C = \frac{1}{2}\left(\sum_{i=1}^{n}\sqrt{A_i D_i v_i}\right)^2. \tag{9}$$

If we define new variables

$$t = x/\sqrt{C} \tag{10}$$

and

$$u = y/\sqrt{C}, \tag{11}$$

we have

$$tu = 1 \tag{12}$$

and the coordinates of the external point are

$$a = x_o/\sqrt{C}, \qquad b = y_o/\sqrt{C}. \tag{13}$$

Note that because the point (x_o, y_o) must be above the hyperbola, equations (12) and (13) imply

$$ab > 1. \tag{14}$$

We use a and b, instead of t_o and u_o, to avoid subscripts in what follows. Figure 2 shows the normalized exchange curve. Using (12), the square of the Euclidean distance from (a,b) to an arbitrary point on the curve with horizontal value t is

$$z(t) = (a-t)^2 + (b-1/t)^2. \tag{15}$$

A necessary (but not necessarily sufficient) condition for the t value that minimizes z (and its square root, which is the distance) is obtained by setting

$$\frac{dz(t)}{dt} = 0.$$

As shown in part 1 of Appendix A, this leads to the following 4th-order equation in t:

$$t^4 - at^3 + bt - 1 = 0 \tag{16}$$

(An alternate way to derive (16) is to use the fact that to achieve the minimum distance the line connecting the external operating point to the point S on the curve must be perpendicular to the tangent of the curve at S. In other words, the product of the slopes of these two lines must be equal to minus 1.)

As will be demonstrated shortly, the behaviour is symmetric in a and b. Thus we focus attention on the situation where $b > a$. Now, a 4th-order equation can have, in general, up to four real roots. However, we are only interested in roots between $1/b$ and a (because, as seen in Figure 2, the minimum distance to the curve from point (a,b) can never be to a point on the curve outside this range). If we denote the left side of (16) by $F(t)$, then in part 2 of Appendix A we prove that $F(1/b) < 0$ and $F(a) > 0$. Thus between $1/b$ and a there has to be precisely either 1 value or 3 values of t where $F(t) = 0$. Extensive numerical calculations (determination of the root(s) and evaluation of $F(t)$ for t values between $1/b$ and a for many a,b combinations with $b > a$ and satisfying (14)) were carried out. Examination of the number of roots and the associated behaviour of the distance as a function of t revealed the following important properties:

i) As shown in Figure 3, the behaviour of the number of roots is symmetric in a and b (we actually knew this through analytic reasoning, as will be described shortly) with 3 roots being in a fan-shaped region starting at the point (2,2).

ii) When there is a single root, it is indeed the value of t (position of S) in Figure 2 where the distance from the external point is minimized.

iii) When there are three roots, the distance as a function of t has two local minima separated by a local maximum, as shown in Figure 4. Moreover, as illustrated in Figure 4, the global minimum of the distance occurs at the smallest of the three roots.

If $a > b$, the whole derivation can be based on using the vertical dimension u, instead of the horizontal dimension t, of point S. Then, instead of (16), we end up with

$$u^4 - bu^3 + au - 1 = 0 \tag{17}$$

and the coefficient of u in (17) is larger in absolute value than that of u^3 (as was the case for t versus t^3 in (16) when $b > a$). Thus for $a > b$ we would determine the single or minimum root of u between $1/a$ and b. It is seen that (16) and (17) are symmetric in a and b as are the limits on t and u. This is the underlying reason for the symmetric behaviour of the number of roots observed in Figure 3.

Because of the symmetry of the results, henceforth we focus the discussion on just the case where $b > a$. Under such circumstances we can summarize as follows. We either have to find the single root or the smallest of the 3 roots of (16) lying between $1/b$ and a. But this is equivalent to simply finding the smallest root between $1/b$ and a. The roots of a quartic equation can be expressed in a complicated algebraic fashion using the approach of Cardano (Boyer 1991), available in Mathematica (Wolfram Research, Inc., 2002) and shown in Appendix B. When there is only one root between $1/a$ and b, there are two imaginary roots.

4. Summary of the procedure for finding the minimum distance and associated point on the exchange curve

Step 1: Normalize the original operating point using (9) and (13) to obtain (a,b). (The rest of the procedure assumes $b \geqslant a$. As discussed in the previous section, if $a \leqslant b$, we are in the southeast portion of Figure 3 and would use the procedure to find the minimum root of the vertical coordinate u of point S - see Figure 2).

Step 2: Find the minimum root of (16) using Cardano's formulae (Appendix B). Denote this root by t^*.

Step 3: Use (15) to get the normalized squared distance $z(t^*)$. Convert back to the original problem dimensions using (10), (11), (12) and

$$Distance = \sqrt{Cz(t^*)} \tag{18}$$

where the latter comes from use of (10), (11), (13) and (15).

5. Illustrative results

For illustrative purposes we use an exchange curve adapted from a case study reported in Peterson and Silver (1979), based on consulting experiences of those authors. For a population of items the C value from (9) was approximately 1,302,400,000 and the estimated current operating point was TFCR=38,400 and TACS=80,000. The hyperbola and the operating point are shown in Figure 5.

We use the procedure of the preceding section to find the minimum distance from the operating point to the exchange curve:

Step 1: From (13)

$$a = 38,400/\sqrt{1,302,400,000} = 1.0640$$

and

$$b = 80,000/\sqrt{1,302,400,000} = 2.2168.$$

Step 2: For this example the four roots of (16) are shown in Appendix B. The only root in the range $1/b$ to a is $t^* = 0.480285$.

Step 3: From (15), $z(t^*) = .35891$. Then (10) to (12) give the coordinates of the estimated point on the curve where the distance is minimized as TFCR = 17,333 and TACS = 75,140. Finally, use of (18) results in an estimated minimum distance of 21,620.

For demonstration purposes, by adjusting the position of the external operating point we were able to generate a variety of (a,b) values. For four illustrative cases the (hypothetical) operating points are shown in Figure 5 and the associated results are presented in Table 1. Note that in two of the cases there are 3 real roots between $1/b$ and a.

6. Conclusions

In this paper we have been concerned with moving from a current operating point to a specified, two-dimensional, efficient frontier. This included a discussion of the managerial relevance of selecting the point on the frontier that is at the minimum distance from the external point. Attention was focussed on a particular efficient frontier relevant in inventory management where the exterior point corresponds to the current or a proposed ordering policy. The frontier in an EOQ context is an hyperbola representing aggregate average cycle stock plotted against total fixed costs of replenishments per year, expressed in the same monetary units. It was shown that finding the minimum distance requires determining the smallest root of a quartic equation within a specified domain.

Our approach is obviously restricted to the situation where the frontier can be represented as an hyperbola. Perhaps there are other managerial tradeoffs (besides the inventory one) where such is the case. Possible extensions of the work would be to deal with other analytic forms (such as a quadratic or parabola) given that these provide accurate representations of tradeoffs of interest to management. As pointed out by Erlenkotter (1989), the inventory cost function of (1) is a special case $(m = 1, n = 1)$ of a more general function

$$f(Q) = C_1 Q^m + C_2 Q^{-n}$$

(with C_1, C_2, m and n all positive). He also indicates that the case of $m = 0.5$ and $n = 1$ occurs in a market area model. Another extension, mentioned in the Introduction, would be towards incorporating the differing revenues and costs associated with moves to different points on the frontier.

Acknowledgements

The research leading to this paper was supported by the Natural Sciences and Engineering Research Council of Canada under Grants A1485, 239153, and 283134-04. The authors express their thanks to an anonymous reviewer and the associate editor, Ton de Kok, for suggestions that markedly improved the original version of the paper.

Appendix A. Derivations

1. *Derivation of (16)*

$$\frac{dz(t)}{dt} = -2(a-t) + 2(b-1/t)(1/t^2) = 0.$$

Multiplying through by t^3 leads directly to (16).

2. *$F(1/b), F(1)$ and $F(a)$ (for $b > a$)*

From (16)

$$\begin{aligned} F(1/b) &= (1/b)^4 - a(1/b)^3 + b(1/b) - 1 \\ &= (1/b)^3(1/b - a) < 0. \\ F(1) &= 1 - a + b - 1 > 0. \\ F(a) &= a^4 - a^4 + ba - 1 \\ &= b(a - 1/b) > 0. \end{aligned}$$

Appendix B. Cardano's Formulae to Find the Roots of (16)

Using Mathematica, the roots of the quartic equation given in (16) are found to be

$$t = \frac{a}{4} - \frac{f}{2} \pm \frac{1}{2}\sqrt{\frac{-(a^3 - 8b)}{4f} + g}$$

and

$$t = \frac{a}{4} + \frac{f}{2} \pm \frac{1}{2}\sqrt{\frac{a^3 - 8b}{4f} + g}$$

where

$$f = \sqrt{\frac{a^2}{4} + \frac{ab - 4}{h} + \frac{h}{3}}$$

$$g = \frac{a^2}{2} - \frac{ab - 4}{h} - \frac{h}{3}$$

and

$$h = \left(\left(27(b^2 - a^2) + \sqrt{-4(3ab - 12)^3 + (27(b^2 - a^2))^2}\right)/2\right)^{\frac{1}{3}}.$$

For the case of $a = 1.0640$ and $b = 2.2168$ given in Section 5 above, these expressions evaluate to $t = -1.17181$, $t = 0.480285$, $t = 0.877784 - 1.00315i$, and $t = 0.877784 + 1.00315i$.

REFERENCES

ALEXANDER,C.A., BUSCH,G., STRINGER,K. (2003) Implementing and interpreting a data envelope analysis model to assess the efficiency of health systems in developing countries. *IMA Journal of Management Mathematics*, **14**, 49-63.

BROWN, R.G. (1967) *Decision Rules for Inventory Management.* New York: Holt, Rinehart and Winston.

BOYER, C.B. (1991) *A History of Mathematics.* 2nd ed., revised by U.C. Merzbach. New York: Wiley.

CARLSON, R. & THERP, H. (1977) A multicriteria approach to strategic planning: an application in inventory control. *OMEGA*, **5**, 57-65.

CHEN, Y. & ZHU, J. (2003) DEA models for identifying critical performance measures. *Annals of Operations Research*, **124**, 225-244.

DOMBERGER, S., FERNANDEZ, P. & FIEBIG, D.G. (2000) Modelling the price, performance and contract characteristics of IT outsourcing. *Journal of Information Technology*, **15**, 107-118.

EATON, J.A. (1964) New - the LIMIT technique. *Modern Materials Handling*, **19**, 38-43.

ERLENKOTTER, D. (1989) The general optimal market area model. *Annals of Operations Research* **18**, 45-70.

GARDNER, E. & DANNENBRING, D. (1979) Using optimal policy surfaces to analyze aggregate inventory tradeoffs. *Management Science*, **25**, 709-720.

HADLEY, G. & WHITIN, T.M. (1963) *Analysis of Inventory Systems*. Englewood Cliffs, N.J.: Prentice Hall.

NOBLE, M.A. (1995) Manufacturing strategy: Testing the cumulative model in a multiple country context. *Decision Sciences*, **26**, 693-721.

PETERSON, R. & SILVER, E.A. (1979) *Decision Systems for Inventory Management and Production Planning.* New York: Wiley.

SCHMENNER, R.W. & SWINK, M.L. (1998) On theory in operations management. *Journal of Operations Management*, **17**, 97-113.

SILVER, E.A., PYKE, D.F. & PETERSON, R. (1998) *Inventory Management and Production Planning and Scheduling.* New York: Wiley.

STARR, M.K. & MILLER, D.W. (1962) *Inventory Control: Theory and Practice.* Englewood Cliffs, N.J.: Prentice Hall.

WAGNER, H.M. (1975) *Principles of Operations Research.* 2nd ed. Englewood Cliffs, N.J.: Prentice Hall.

WOLFRAM RESEARCH, INC. (2002) *Mathematica*, Version 4.2. Champaign, IL.

List of Captions and Titles

Figures

Tables

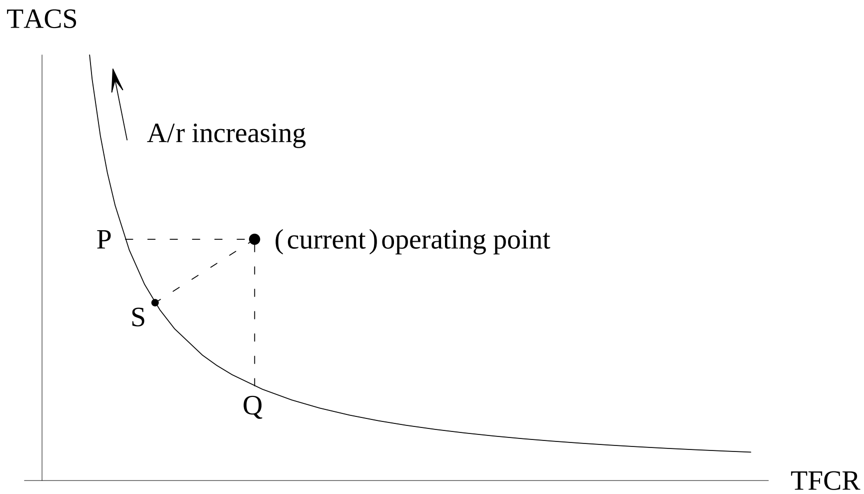

Figure 1: Generic representation of the EOQ exchange curve

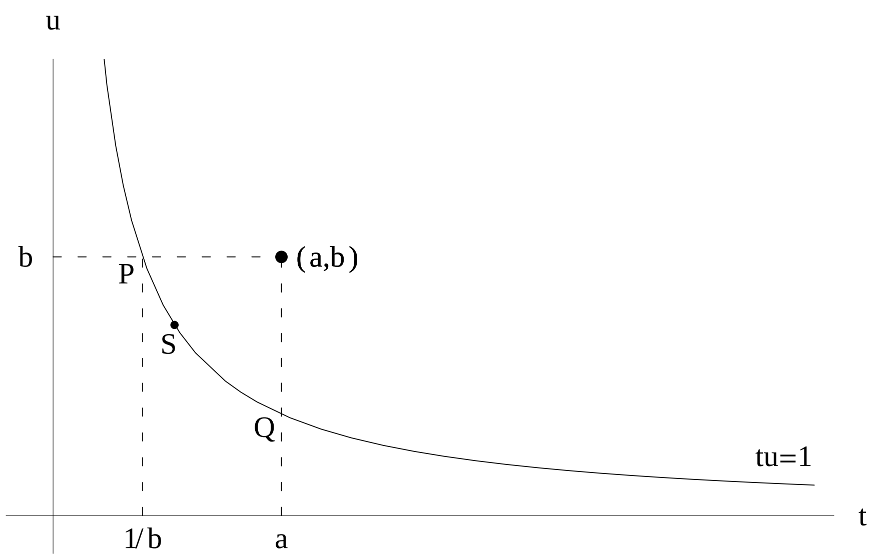

Figure 2: Normalized exchange curve

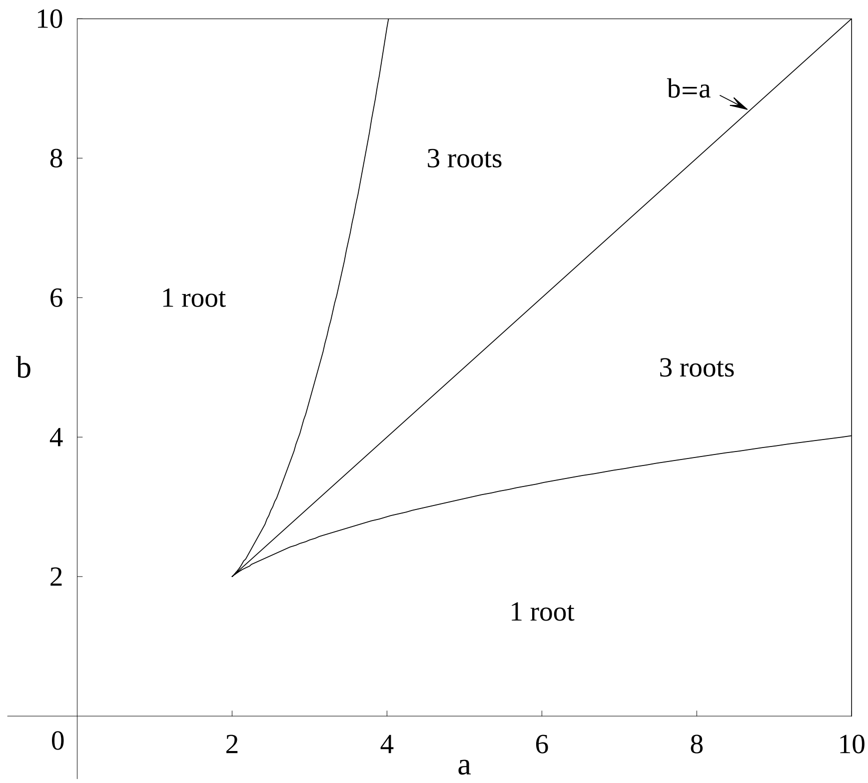

Figure 3: Number of roots as a function of *a* and *b* (i.e., the location of the external operating point)

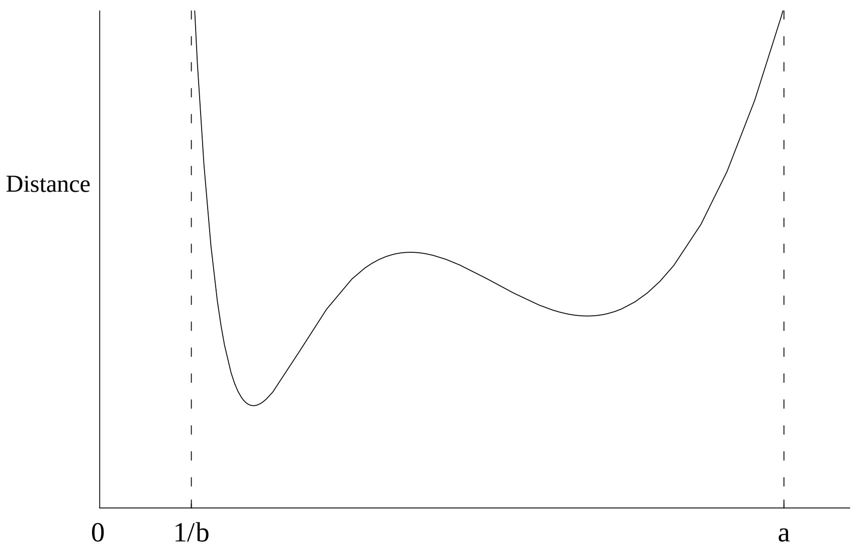

Figure 4: Distance function for $b > a$ when there are 3 roots of (16)

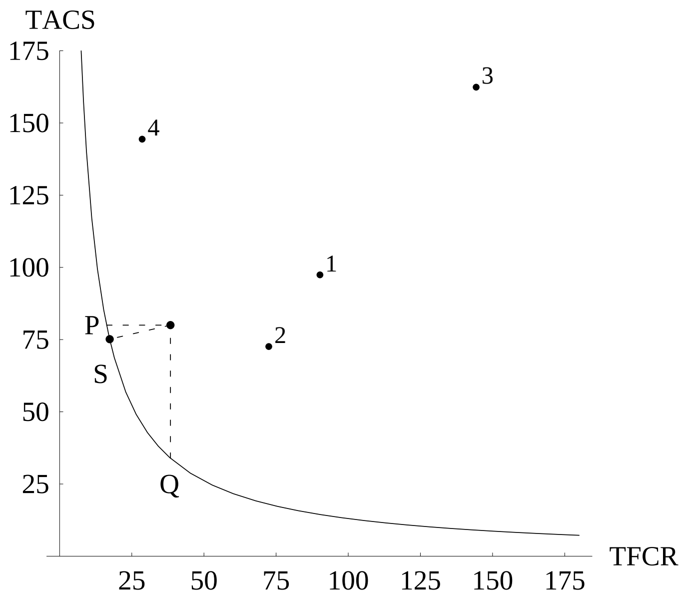

Figure 5: Illustrative examples (units are in thousands)

Example Number	TFCR	TACS	a	b	Number of Roots Between $1/b$ and a	t^*	Closest Point on Exchange Curve TFCR	TACS
1	90,200	97,400	2.4994	2.6989	3	.43246	15607	83449
2	72,500	72,600	2.0089	2.0117	1	.86776	31316	41589
3	144,300	162,400	3.9985	4.5000	3	.23278	8400	155037
4	28,600	144,400	0.7925	4.0012	1	.25209	9097	143161

Table 1: Additional illustrative results