The Fixed Rate Equivalent (FREQ): Measuring the Performance of Financial Accounts in the Presence of Deposits and Withdrawals

A GreaterThanZero White Paper

July 7, 2012

Abstract

A common way of measuring the performance of a financial account in a way that takes into account deposits and withdrawals (money-weighted performance) is to calculate an internal rate of return (IRR). This amounts to modeling the account's performance with that of a fixed rate account: the IRR can be interpreted as the fixed interest rate that replicates the performance of the account. However, from the point of view of an individual investor, the IRR's model is unrealistic: the fixed rate account by which the IRR models the investor's account has the property that it charges the same interest rate on a negative balance that it pays on a positive balance. Teichroew et al. have shown that this way of charging interest is also the root cause of the IRR's mulitple solution problem, and that a more realistic way of charging interest on negative balances results in unique solutions (see [1] and [2]). The articles by Teichroew et al. are written in the context and for the purpose of capital budgeting. The purpose of this paper is to explain and prove all this in a manner that is geared towards measuring the performance of an individual investor's accounts. We call the rate that results from replacing the IRR's fixed rate model with a more realistic one the *fixed rate equivalent*, or *FREQ* for short. We also show how the FREQ can be efficiently calculated.

1 Introduction

The *internal rate of return* (IRR), also known as the *discounted cash flow rate of return* (DCFROR), or the *effective rate of return*, is a measure of profitability that is used in capital budgeting as well as in the world of individual investing, where it sometimes appears under the moniker *personalized rate of return*. The IRR deals with situations where a sequence of cash flows to and from an investment occur at certain points in time. The IRR is most commonly defined as a real number greater than -1 that solves the equation

$$\sum_{i=1}^{m} \frac{c_i}{(1+x)^{p_i}} = 0 \tag{1}$$

where

m = number of cash flows

c_i = amount of the i^{th} cash flow

p_i = number of periods between the first and the i^{th} cash flow (hence $p_1 = 0$).

In plain language, this definition says,"The internal rate of return is the rate that makes the sum of the net present values of the cash flows equal to zero, where 'present' refers to the time of the first cash flow."

As we will explain in more detail below, one can rearrange Equation (1) in such a way that it reveals the following alternate interpretation of the internal rate of return: "The IRR is the interest rate of the hypothetical fixed-rate investment which, when subjected to the same cash flows as the real investment, results in the same gain or loss as the real investment." For ease of terminology, we will also use the term *replication* to describe this situation: the IRR of an investment is the fixed interest rate that *replicates* the investment. This interpretation has the added side benefit that it is intuitive for the investor who does not have a background in mathematics or economics: "Had I put my money in a savings account, this is the rate that would have gotten me to where I am now."

Upon closer inspection, however, there is an issue with the internal rate of return under this point of view. The fixed-rate investment that the IRR uses to replicate the actual investment has the property that it always applies the same rate, even if the balance becomes negative. That is not how real-life fixed-rate investments behave: the rate that a financial institution charges on a negative balance is substantially different from the one that it pays on a positive balance. Moreover, the IRR may well be negative. In that case, the IRR's model amounts to paying an investor for being in debt. This is all the more troubling because—as we will demonstrate later—the balance of the modeling fixed rate account can become negative even if the investor's account did not.

In [1] and [2], Teichroew et al. have shown that charging the same interest rate on positive and negative balances is also the root cause of the well-known fact that Equation (1) may have more than one solution. This is a common objection to the use of the IRR in the world of individual investing: how can an advisor discuss the performance of an account with a client when, according to the IRR, that performance is 5% as well as 8%? How does one compare the performance of two accounts when the performance of one is 5% and 8%, and the performance of the other is 4% and 9%?

In view of this situation, it seems natural to measure the performance of financial accounts by a rate that is modelled after the IRR, except that the replicating fixed rate account charges the actual, realistic cost of borrowing on negative balances.

Definition 1.1 Suppose that we are given the following information about an individual investor's financial account:

(i) A begin date and balance.

(ii) An end date and balance.

(iii) The dates and amounts of all deposits and withdrawals in between.

Then we define the **fixed rate equivalent**, or **FREQ** for short, of the account for that time period as the annual interest rate of the fixed rate account which, when seeded with the same beginning balance and subjected to the same deposits and withdrawals as the real account, results in the same ending balance as the real account. Here, the fixed rate account is such that in case of negative balances, the actual cost of borrowing is charged. □

As a measure of performance for an individual investor's financial accounts, the FREQ has the following advantages:

(i) Grasping the intuitive meaning of the FREQ does not require any mathematical understanding at all. It can be explained to the investor by saying, "Had you put your money in a savings account, this is the rate that would have taken you to the same ending balance."

(ii) According to [1] and [2], the FREQ does not suffer from the multiple solutions problem that plagues the IRR. Moreover, a unique FREQ exists except in certain corner cases, such as the investor always being in debt.

(iii) The FREQ is consistent with the ordinary annualized return as defined in modern portfolio theory: in the absence of any deposits and withdrawals, the FREQ equals the annualized return.

The purpose of this paper is to give a complete and self-contained presentation of the FREQ and the mathematics that surrounds it. Section 2 provides explanations and illustrations that are mathematically rigorous but skip all proofs. Sections 3 and 4 are pure mathematics, providing the missing proofs.

Let us emphasize again that there are no new mathematical results in this paper. The core mathematics is already contained in [1] and [2]. However, [1] and [2] are written in the context of capital budgeting and financing decisions. Applying the results to the performance measurement of investor accounts with deposits and withdrawals requires some adapting. The purpose of this paper is to provide a presentation that saves the reader the trouble of having to perform that adaptation.

2 From the IRR to the FREQ

Since the FREQ is a modification of the IRR, we begin by taking a closer look at the IRR. In order to do so, we need to address two trivial but potentially confusing issues. Firstly, in the world of capital budgeting, cash flows into an investment are usually given as negative numbers (money out of the investor's pocket), while cash flows out of the investment are given as positive numbers (money into the investor's pocket). In the world of individual investing, it is the other way round. On a brokerage statement, deposits into an account are positive numbers and withdrawals are negative numbers. *Throughout this paper, we will use the individual investors' convention, counting inflows to the investment as positive numbers, outflows as negative numbers.* Needless to say, switching between the two points of view amounts to no more than multiplying both sides of Equation (1) by a factor of -1, which has no bearing on the solvability and solutions of the equation.

The second issue that needs clarification is even more trivial than the first one. It is clear that only real numbers greater than or equal to -1 make sense as rates of return. Moreover, the classical definition of the IRR as shown in Equation (1) above implicitly excludes the number -1 as a solution by putting the expression $1 + x$ in the denominator. We will therefore use the following terminology throughout this paper:

Definition 2.1 By a **rate**, we mean a real number greater than -1. Specifically, a solution to Equation (1) that is greater than -1 will be called a **solving rate** for the equation.

In Equation (1) above, we have already stated the most common definition of the IRR (see e.g. [4]). The first step towards the FREQ is to transform Equation (1) in such a way that it reveals the alternate interpretation of the IRR as the interest rate of the fixed-rate investment that replicates the actual investment. This transformation is mathematically trivial. However, in order to keep track of what the individual pieces of the equation mean, we have to perform it in several small steps.

If we multiply both sides of Equation (1) by $(1 + x)^{p_m}$ and then move the constant coefficient to the right hand side, we obtain

$$\sum_{i=1}^{m-1} c_i \cdot (1+x)^{p_m - p_i} = -c_m. \tag{2}$$

To make this equation look simpler, let us introduce the following new notation:

$n = m - 1$, the number of cash flows except for the last

$t_i = p_m - p_i$, the number of periods between the i^{th} and the last cash flow

$c = -c_m$, the negative of the last cash flow

Equation (2) now becomes

$$\sum_{i=1}^{n} c_i \cdot (1+x)^{t_i} = c. \tag{3}$$

Finally, we successively pull the highest possible power of $(1+x)$ out of the leftmost two summands in Equation (3):

$$(\ldots((c_1 \cdot (1+x)^{t_1-t_2} + c_2) \cdot (1+x)^{t_2-t_3} + c_3) \cdot \ldots + c_n) \cdot (1+x)^{t_n} = c \tag{4}$$

If one wanted to evaluate the left hand side of Equation (4) for a particular value r of x, one would do it from the inside out, successively evaluating the following expressions:

$$\begin{array}{rcl} H_1(r) & = & c_1 \\ H_2(r) & = & H_1(r) \cdot (1+r)^{t_1-t_2} + c_2 \\ H_3(r) & = & H_2(r) \cdot (1+r)^{t_2-t_3} + c_3 \\ & \vdots & \\ H_n(r) & = & H_{n-1}(r) \cdot (1+r)^{t_{n-1}-t_n} + c_n \\ H_{n+1}(r) & = & H_n(r) \cdot (1+r)^{t_n} \end{array} \tag{5}$$

This way of evaluating a polynomial is known as the Horner scheme, and it is in fact used in mathematical software as the most efficient way of evaluating a polynomial. In the case of the internal rate of return, it reveals the alternate interpretation of the IRR that we were looking for: if we interpret the first cash flow as an initial deposit, then the left hand side of Equation (4) is the ending balance of an account with a fixed interest rate of x, subjected to the given deposits and withdrawals except for the last one. If we also interpret the last cash flow as the withdrawal that empties the account (or, equivalently, as the negative of the ending account balance), then solving the equation for x means finding the fixed interest rate that leads to the same ending balance as the actual investment. In other words, the internal rate of return aims to be the fixed interest rate that replicates the actual investment. Equations (5) above demonstrate this nicely: the expressions

$$H_i(r) \quad (1 \leq i \leq n)$$

are the account balances right after the i-th cash flow, and $H_{n+1}(r)$ is the ending balance.

Let us consider an example. Suppose the following activities occur in an investor's account:

Table 1

Date	1/1/1992	1/1/1995	1/1/2000	1/1/2002
Cash Flow	$5,000	-$10,000	$6,000	-$672.08

Recall that the \$5,000 on 1/1/1992 stand for the initial deposit, and the final withdrawal of \$672.08 empties the account, meaning that \$672.08 was the ending balance on 1/1/2002. The scenario of Table 1 is a rather realistic one for an individual investor. Someone could have opened a brokerage account in 1992 with \$5,000, more than doubled the account balance during the bull market of the 1990's, and then taken \$10,000 of gains off the table in 1995. After another fairly large deposit in January of 2000, the investor ended up losing almost everything in the crash of the spring of 2000.

If you enter the numbers of Table 1 into your favorite software for calculating the IRR, then, depending on which software you use and what initial guess you enter (if any), the answer could be 5.5% or 8.66%. We are looking at a case where there is more than one IRR.

As we have mentioned before, the root cause of the IRR's multiple solutions issue is the fact that the IRR uses a model where the fixed rate account charges the same rate regardless of whether the balance is positive or negative. The point of the present example is to illustrate that phenomenon. Here's the equivalent of Equation (4) for our present example:

$$((5000 \cdot (1+x)^3 - 10000) \cdot (1+x)^5 + 6000) \cdot (1+x)^2 = 672.08 \tag{6}$$

First of all, one can use this to verify the two IRRs of 5.5% and 8.66%. (In fact, one must use a little more precision to get the end value right to the penny; use 8.663 instead of 8.66.) Recall that the left hand side of Equation (6) describes how the account balance of a savings account with a fixed annual interest rate of x develops over time if we make the deposits and withdrawals of Table 1. Now let us use the equivalent of expressions (5) to follow the balance of that fixed rate account over time. We'll do this for the 5.5% interest rate. For the first three years, we have \$5,000 compounding at 5.5%, followed by a \$10,000 withdrawal. The hypothetical account balance on Jan 1, 1995 is thus

$$5000 \cdot (1+.055)^3 - 10000 = 5871.21 - 10000 = -4128.79 \tag{7}$$

For the next five years, we are looking at -\$4128.79 compounding at 5.5%, followed by a \$6,000 deposit into the account. Therefore, the calculation of the account balance for Jan 1, 2000 looks like this:

$$-4128.79 \cdot (1+.055)^5 + 6000 = -5396.17 + 6000 = 603.83 \tag{8}$$

We see that as the IRR tries to model the actual investment as an account with a fixed rate, the balance of that fixed-rate account drops into the negative. The model used by the IRR then applies ordinary compounding to that negative balance. That amounts to charging the same interest rate on a debt that is paid on a positive balance. Note that in the case of a negative IRR, this actually amounts to paying an investor for being in debt. The following theorem is a mathematically precise statement of the fact that the use of this particular fixed-rate model is the root cause of the IRR's multiple solutions issue.

Theorem 2.2 Let r be a rate that solves Equation (4). Suppose that r has the following property: the expressions

$$H_i(r) \quad (1 \leq i \leq n)$$

as defined in (5) are all non-negative, that is, the balance of the fixed-rate account by which the IRR models the investment never becomes negative. Then r is the only rate that solves the equation.

A proof of Theorem 2.2 is given in Section 3 of this paper, where it appears as Theorem 3.5. Let us emphasize again that it is not necessary to study our proof of Theorem 2.2 for mathematical verification. The core mathematics is already contained in the papers by Teichroew et al. ([1] and [2]). Our presentation is merely structured differently, according to our purpose of adapting the results of Teichroew et al. to the problem of measuring the performance of investor accounts with deposits and withdrawals (money-weighted performance).

There is a subtle strength in Theorem 2.2 that is easy to overlook. A casual reader might misinterpret the statement of the theorem as follows: for a given investment, there can be at most one internal rate of return for which the balance of the replicating fixed-rate account never becomes negative. However, the theorem is much stronger than that. It says that if there is an IRR for which the balance of the replicating fixed-rate investment never dips into the red, then there can be no other IRR at all.

In connection with the mathematical proof of Theorem 2.2, we will also prove in Section 3 below that the converse of the theorem is not true: if we have an investment for which the IRR is unique, then it is not necessarily true that the balance of the replicating fixed-rate account is never negative. Therefore, hunting for mathematical conditions that guarantee uniqueness of the IRR is not, in our opinion, the best way to resolve the multiple solutions issue. Instead, one should modify the model by which the fixed rate is calculated on positive and negative balances (see Theorem 2.3 below).

The mathematical proof of Theorem 2.2 is not trivial. However, the example of Table 1 can be used to gain an intuition for why and how a negative balance in the fixed-rate account that replicates the investment creates the multiple solutions issue for the IRR. Recall that in this example, the IRR has the two values 5.5% and 8.66%. From Theorem 2.2, we already know that for neither one of the two rates, the balance of the replicating fixed-rate account stays non-negative all the way. In both cases, there is compounding on a negative balance going on. Otherwise, there could not have been more than one solving rate.

Let us now look at the account balances over time in the fixed-rate accounts for both rates. We've already calculated two of these balances in (7) and (8) above. Continuing in the same way, we obtain the following table of account balances (the balance on each date is the balance after the deposit or withdrawal on that date has been processed, with the exception of the last date, where we show the balance just before the $672.08 withdrawal empties the account):

Table 2

Date	1/1/1992	1/1/1995	1/1/2000	1/1/2002
Account Balance at 5.5%	$5,000	-$4,128.79	$603.83	$672.08
Account Balance at 8.66%	$5,000	-$3,584.73	$569.19	$672.08

Here is a chart for the data of Table 2:

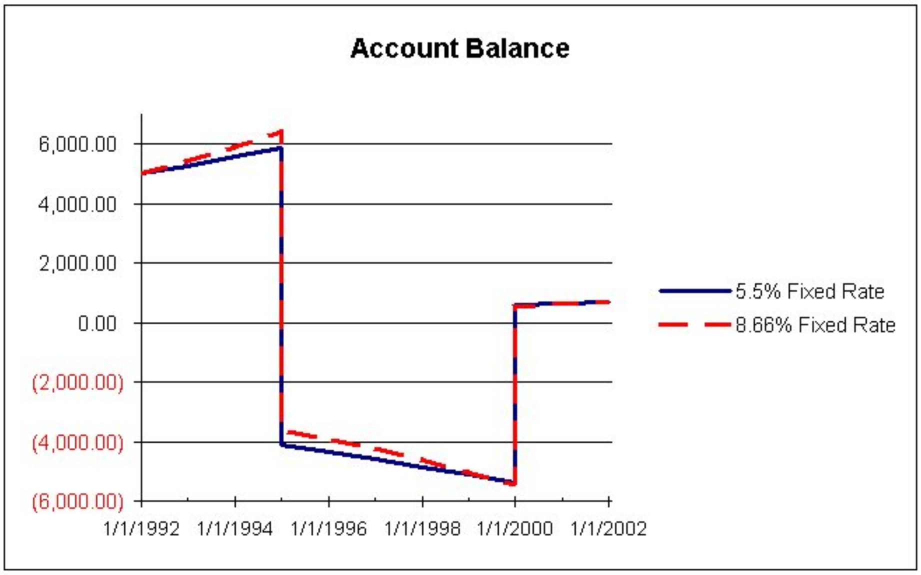

Chart 1

Imagine that you start with the 5.5% graph, then increase the interest rate so as to morph it into the 8.66% graph. The positive account values during the first three years will increase faster, while the account values during the subsequent five years will decrease faster. That's how it is possible for the ending balance to be the same for the two rates. The intuition behind Theorem 2.2 is that this kind of "opposite movement" cannot happen unless the account balance dips into the negative.

So far we have seen:

(i) The internal rate of return, when interpreted as the interest rate paid by the fixed-rate investment that replicates the actual investment, uses a model that applies the same compounding rate regardless of whether the balance is positive or negative. In the world of individual investing, this is unlikely to be deemed an appropriate model.

(ii) This model of applying interest is the root cause of the multiple solutions issue of the IRR.

How can this situation be amended? There is really only one acceptable way of modifying the model of the fixed-rate investment: do what is done in real life, that is, charge the cost of borrowing that was in effect at the respective time period whenever the balance of the modeling fixed-rate investment becomes negative. The following theorem says that this model leads to a rate of return that no longer suffers from the multiple solutions issue. For simplicity, the theorem will assume that the first cash flow is positive (initially positive account balance) and the last cash flow is negative (ending account balance positive). This will exclude some odd corner cases such as the one where the investor is always in the red. The mathematically precise version of the theorem that is given in Section 4 below uses more refined assumptions.

Theorem 2.3 Suppose we are given a series of cash flows as in Equation (1), and the first cash flow is positive, while the last one is negative. Assume further that for each period of time between two cash flows, we know what the cost of borrowing is, i.e., the interest rate charged on an account with a negative balance. Then there is exactly one rate r with the following property: the fixed-rate investment whose interest rate on a positive balance equals r and whose rate charged on a negative balance is the given cost of borrowing replicates the given investment. Moreover, there is an algorithm that calculates the rate r from the given cash flows and the cost of borrowing.

Definition 2.4 We call the rate described above the **fixed rate equivalent**, or **FREQ** for short, for the given investment and cost of borrowing.

The mathematical proof of Theorem 2.3 is given in Section 4 of this paper, where it appears as Theorems 4.8 and 4.10. The assumptions about the first and last cash flow will be replaced with a more refined hypothesis. As with Theorem 2.2, it is not necessary to study our proof of Theorem 2.3 for mathematical verification. We are merely giving a differently structured presentation of the results of [1] and [2], geared toward the purpose of measuring the performance of investor accounts with deposits and withdrawals (money-weighted performance).

At first glance, it appears as if the calculation of the FREQ would be much harder than that of the IRR because of the need to provide the cost of borrowing. In practice, this turns out to be less of a burden than it would seem. What one should do is to start with some rough estimate of the cost of borrowing. In fact, that estimate could be anything at all; accuracy does not matter. Next, one calculates the FREQ for that cost of borrowing. For real-life investments, it is quite likely that the balance of the replicating fixed-rate investment never goes negative, so that the cost of borrowing never applies. That means it would not have applied for any other estimate either: the cost of borrowing does not matter for this investment. The unique FREQ has been found, and by Theorem 2.2, the FREQ then also equals the unique IRR.

If, on the other hand, the FREQ calculation reveals that the replicating fixed rate investment did become negative, that is, the cost of borrowing was applied at some point, things get subtle for performance measurement. Suppose we go back and supply the real-life cost of borrowing. (Recall that the cost of borrowing used may vary over time as is the case in real life.) With that cost of borrowing, we may then calculate a realistic FREQ: the annual interest rate of the fixed rate account that replicates the investment's performance and charges the actual cost of borrowing on negative account balances. The problem is that this information, the FREQ and the cost of borrowing, still does not capture the performance very well: much depends on how much time the replicating fixed rate account spent in positive and negative territory, respectively, and what the magnitude of the positive and negative balances were. One may well argue that the best one can do in this situation is to look at a chart of the account balance over time of the fixed rate account and thereby get a visual impression of the performance (see the Chart 2 below for an example). However, it is important to note that one may still use the FREQ for comparing the performance of investments with the same deposits and withdrawals: a higher FREQ is always better than a lower FREQ, assuming that the same cost of borrowing is used.

There is a second option for dealing with the case where the replicating fixed rate account encounters negative balances. In the example that we've been using throughout, the negative balance of the replicating fixed rate account is caused by the fact that a period of very good performance was followed by a large withdrawal, which was then followed by a period of poor performance. This kind of extreme behavior is typical of investments where the replicating fixed rate account goes negative. Therefore, one may, if it is deemed acceptable, try to split the analysis period up into two or more subperiods of more homogeneous performance. There is a good chance that this will once again lead to FREQ calculations that do not use the cost of borrowing.

Because of the fact that the replicating fixed-rate account used by the FREQ rarely becomes negative, there is not much visible difference between the FREQ and the IRR: for many real-life investments, the IRR is unique and equals the FREQ. However, this does not mean that the FREQ provides little or no gain in practice. The IRR is typically calculated using numerical methods such as Newton's method. The investor is then told that this is "the" IRR. However, there is no information as to whether this IRR value is unique, or, more importantly, whether the underlying replicating fixed-rate investment did or did not encounter negative balances. The FREQ, on the other hand, always gives the guarantee of uniqueness, and, for that matter, of having used a realistic and intuitive model for the fixed-rate replication. Loosely speaking, one could make the case for the FREQ like this: the FREQ is unlikely to change your IRR, but it is the only way of making sure that your IRR is good.

To illustrate the FREQ and give an intuition for why it does not suffer from the multiple solutions issue that plagues the IRR, let us return to the example of Table 1. Recall that in Chart 1, we have shown how and why there are two IRRs in that case. Now let us assume a cost of borrowing of 20% for the period 1/1/1995 through 12/31/1996, and 5% for the period 1/1/1997 through 12/31/1999. (We chose these somewhat extreme values for the sake of better visualization.) The FREQ, that is, the interest rate that replicates the investment under these assumptions, equals 10.41907%. In the chart below, we also show the account balances over time for accounts with a 12% interest rate and a 9% interest rate, respectively. This demonstrates that the "opposite movement" that lead to the multiple IRR values as shown in the Chart 1 cannot happen anymore when the FREQ is used: a higher interest rate will lead to a higher ending balance, and a lower interest rate will lead to a lower ending balance.

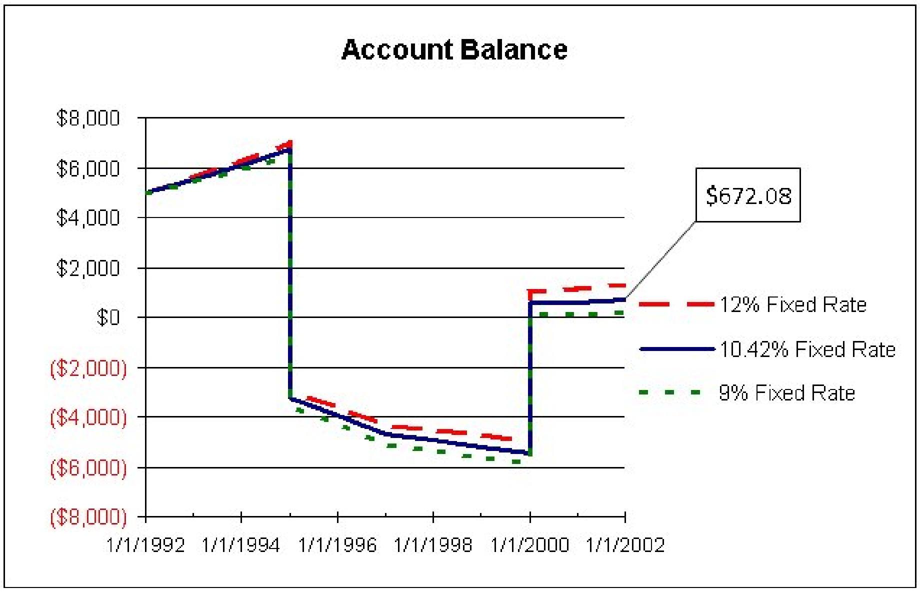

Chart 2

The table below shows the exact balances for the accounts of Chart 2 above.

Table 3

Date	1/1/1992	1/1/1995	1/1/1997	1/1/2000	1/1/2002
Balance at 9%	$5,000	-$2,975.36	-$4,284.52	$1.040.13	$1,305.00
Balance at 10.42%	$5,000	-$3,268.65	-$4,706.85	$551.23	$672.08
Balance at 12%	$5,000	-$3,524.68	-$5,075.79	$124.14	$147.00

3 The Mathematics of Multiple IRRs

The purpose of this section is to prove Theorem 2.2. For the sake of mathematical rigor, the theorem has been rephrased as Theorem 3.5. After studying the notation and definitions below, the reader should not have any trouble seeing that Theorem 2.2 and Theorem 3.5 are the same. The formal differences can be summed up as follows:

(i) The expression $1+x$ of Equation (4) will be replaced with the variable X.

(ii) The notation $c_1, \ldots, c_n$ for the cash flows will be replaced by $a_n, \ldots, a_1$, because that is the default mathematical notation for the coefficients of a univariate polynomial whose constant coefficient is zero.

(iii) The term *admissible solution* will be used for solutions where the account balance of the replicating fixed-rate account does not dip into the negative.

(iv) The definition of the Horner polynomials (see Equations (5)) will be based on the dense rather than the sparse representation of the underlying polynomial.

Let P be a univariate polynomial with real coefficients whose constant coefficient is zero. We are interested in the uniqueness of positive real solutions of the equation

$$P(X) = a \quad \text{with} \quad a \in \mathbb{R}. \tag{9}$$

Before we begin, a few trivial remarks. First off, we could of course move the constant a to the left hand side of the equation, incorporate it in the polynomial P, and just say that we're investigating the uniqueness of positive real zeroes of univariate polynomials. The only reason for phrasing our problem as in (9) above is that it suits our statements and proofs a little better. Secondly, our problem and many of our results make sense not only for the real numbers, but for any ordered field. However, we do not see much benefit from that generalization. Therefore, we restrict our attention to the real numbers. Also, to avoid being overly verbose, we will sometimes use the term"positive solution" as a synonym for "positive real solution."

Rather obviously, positive real solutions of Equation (9) are not unique in general. For example, the equation

$$X^3 - 6X^2 + 11X = 6$$

has the positive solutions $X = 1$, $X = 2$, and $X = 3$. The topic of this section is to investigate how uniqueness of the solution is affected if we focus on solutions that satisfy a certain condition. After some preparation, we will state that condition in Definition 3.3 below.

Definition 3.1 Let

$$P = \sum_{i=0}^{n} a_i X^i$$

be a univariate polynomial whose coefficients are real numbers. Then we recursively define the **Horner sequence for *P*** to be the following sequence of polynomials:

$$\begin{aligned} H_0 &= a_n \\ H_k &= H_{k-1} \cdot X + a_{n-k} \quad \text{for } 1 \leq k \leq n. \end{aligned}$$

The Horner sequence for P is the sequence of expressions that is used in Horner's algorithm (also known as the Horner scheme) for the fast evaluation of polynomials. The correctness of Horner's algorithm rests on the fact that $P = H_n$. This is of course a well-known fact. For easier reference, we give the proof here:

Lemma 3.2 Let P and $H_0, H_1, \ldots, H_n$ be as above. Then $P = H_n$.

Proof We begin by proving an auxiliary claim. Let

$$Q = \sum_{i=0}^{n-1} a_i X^i,$$

and let $K_0, K_1, \ldots, K_{n-1}$ be the Horner sequence for Q. We claim that

$$H_i = a_n \cdot X^i + K_{i-1} \quad \text{for } 1 \leq i \leq n.$$

The proof of the claim is by induction on i. If $i = 1$, then the claim is immediate from the definitions:

$$H_1 = a_n \cdot X + a_{n-1} = a_n \cdot X^1 + K_0.$$

For the induction step, let $i > 1$. The induction hypothesis states that

$$H_{i-1} = a_n \cdot X^{i-1} + K_{i-2}.$$

If follows that

$$H_i = H_{i-1} \cdot X + a_{n-i} = (a_n \cdot X^{i-1} + K_{i-2}) \cdot X + a_{(n-1)-(i-1)} = a_n \cdot X^i + K_{i-1}.$$

We are now in a position to prove the lemma. The proof is by induction on n. For the base case of the induction, let $n = 0$. Then trivially $P = a_n = H_n$. For the induction step, let $n > 0$. As in the claim above, let

$$Q = \sum_{i=0}^{n-1} a_i X^i,$$

and let $K_0, K_1, \ldots, K_{n-1}$ be the Horner sequence for Q. By the induction hypothesis, we have $Q = K_{n-1}$. Using the claim that we proved earlier, we see that

$$P = a_n X^n + Q = a_n X^n + K_{n-1} = H_n. \quad \square$$

From now on, we will be talking about univariate polynomials with zero constant coefficient and their Horner sequences. To avoid repetition, we now state the notation and assumptions that will be in effect for the rest of this section.

Notation and Assumptions
For the rest of this section, let

$$P = \sum_{i=0}^{n} a_i X^i$$

be a non-constant univariate polynomial with real coefficients and zero constant coefficient, that is,

$$a_0, \ldots, a_n \in \mathbb{R}, \quad n > 0, \quad a_0 = 0, \quad \text{and} \quad a_n \neq 0.$$

We will denote the Horner sequence of P by $H_0, H_1, \ldots, H_n$.

We are now in a position to state the condition that we will require of our solutions to Equation (9).

Definition 3.3 With the notation and assumptions stated above, we call the set

$$A(P) = \{\, r \in \mathbb{R} \,|\, r > 0 \text{ and } H_i(r) \geq 0 \text{ for } 0 \leq i \leq n \,\}$$

the **set of admissible values for $\boldsymbol{P}$**. A solution to Equation (9) is called **admissible** if it is an admissible value for the polynomial P on the left hand side of the equation.

Recall from the previous section that in the context of fixed-rate investments, each Horner polynomial $H_i(r)$ represents the balance after i periods of the fixed-rate investment whose cash flows are the coefficients of P. Therefore, in that context, an admissible rate r is a rate for which the balance of the investment never becomes negative. This observation will make it easy to see how Theorem 3.5 is just a more formal version of Theorem 2.2.

The following lemma states that the defining condition of an admissible value needs to be checked only for those Horner polynomials where the corresponding coefficient of P is negative.

Lemma 3.4 With the notation of Definition 3.3 above, we have

$$A(P) = \{\, r \in \mathbb{R} \,|\, r > 0 \text{ and } H_i(r) \geq 0 \text{ for all } i \text{ with } a_{n-i} < 0 \,\}.$$

Proof The inclusion from left to right is trivial. For the inclusion from right to left, let r be an element of the set on the right hand side. We use induction on i to show that

$$H_i(r) \geq 0 \quad \text{for} \quad 0 \leq i \leq n.$$

For the base case of the induction, assume for a contradiction that $H_0(r) < 0$. Since $H_0(r) = a_n$, we get $a_n < 0$. Since r is an element of the set on the right hand side, we may conclude that $H_0(r) \geq 0$, contradicting our assumption that $H_0(r) < 0$.

If $i > 0$, we have

$$H_i(r) = H_{i-1}(r) \cdot r + a_{n-i}.$$

If $a_{n-i} < 0$, the claim is immediate from the definition of the set on the right hand side. Else, it follows from the induction hypothesis together with the fact that $r > 0$. $\square$

The rest of this section is devoted to proving and further exploring the following main theorem, which makes a curiously strong statement about the uniqueness of those solutions to Equation (9) that are admissible for P.

Theorem 3.5 With the notation of Definition 3.3 above, suppose that r is an admissible solution to Equation (9), that is,

$$P(r) = a \quad \text{with} \quad r \in A(P).$$

Then $P(s) \neq a$ for all $s \in \mathbb{R}$ with $s > 0$ and $s \neq r$.

In plain language, the theorem states that if r is an admissible value for P that solves Equation (9), then it is the only positive real solution. An equivalent statement would be, "If there is more than one positive real solution to Equation (9), then none of them is admissible."

A casual reader of the theorem might have rephrased it erroneously as follows: "If we restrict the set of positive real solutions to Equation (9) to admissible values for P, then the solution, if any, becomes unique." According to that statement, there could be a situation where there is an admissible solution and a second positive solution that is not admissible. Restricting to admissible solutions would throw out the inadmissible solution and thus make the solution unique. But the statement of the theorem is much stronger than that. It says that if there is an admissible solution, then there cannot be another positive real solution, admissible or not. In other words, admissible solutions are unique as positive real solutions.

Before delving into the proof of Theorem 3.5, let us look at two simple examples.

Example 3.6 The equation

$$X^3 - 6X^2 + 11X = 6$$

has the positive solutions $X = 1$, $X = 2$, and $X = 3$. According to the theorem, none of these can be an admissible value for the polynomial

$$P = X^3 - 6X^2 + 11X,$$

because if one of them were an admissible value for P, then there could not exist another positive real solution. Indeed, we have $H_1 = X - 6$, which is negative at $X = 1$, $X = 2$, and $X = 3$.

Example 3.7 The equation

$$X^2 + X = 2$$

has the positive solution $X = 1$. By Lemma 3.4, this solution is an admissible value for $P = X^2 + X$ since P has no negative coefficients. By the theorem, there cannot be another positive real solution. Indeed, the only other solution is $X = -2$.

We now prove some results that will be instrumental in the proof of Theorem 3.5.

Lemma 3.8 With the notation of Definition 3.3 above, assume that $r_1, r_2 \in \mathbb{R}$ with $r_1 \in A(P)$ and $r_1 < r_2$. Then

$$H_i(r_1) < H_i(r_2) \quad \text{for} \quad 1 \leq i \leq n.$$

Proof First, note that $0 < r_1$ since admissible values are positive by definition. We prove the claim by induction on i. For $i = 1$, we start with

$$a_n = H_0(r_1) = H_0(r_2).$$

Since $0 \leq H_0(r_1)$ and $a_n \neq 0$, we actually have

$$0 < a_n = H_0(r_1) = H_0(r_2).$$

We conclude, using the assumption $0 < r_1 < r_2$, that

$$H_1(r_1) = H_0(r_1) \cdot r_1 + a_{n-1} < H_0(r_2) \cdot r_2 + a_{n-1} = H_1(r_2).$$

Now let $i > 1$. By the induction hypothesis and the fact that $r_1 \in A(P)$, we have

$$0 \leq H_{i-1}(r_1) < H_{i-1}(r_2).$$

Using the assumption $0 < r_1 < r_2$, we may conclude that

$$H_i(r_1) = H_{i-1}(r_1) \cdot r_1 + a_{n-i} < H_{i-1}(r_2) \cdot r_2 + a_{n-i} = H_i(r_2),$$

as desired. □

Proposition 3.9 If the leading coefficient of P is negative, then the set $A(P)$ of admissible values for P is empty. Otherwise, if the leading coefficient is positive, it is a non-empty, upward closed set, that is, $A(P) \neq \emptyset$ and

$$r_1 \in A(P) \wedge r_1 < r_2 \Longrightarrow r_2 \in A(P) \quad \text{for all} \quad r_1, r_2 \in \mathbb{R}.$$

Proof If the leading coefficient of P is negative, that is, $a_n < 0$, then $H_0(r) = a_n < 0$ for all $r \in \mathbb{R}$, and hence there can be no admissible values for P.

Now assume that $a_n > 0$. To see that the set of admissible values for P is not empty, note that every element of the Horner sequence of P is a polynomial whose leading coefficient is positive, and they are all non-constant except for the first one. Therefore,

$$\lim_{r \to +\infty} H_i(r) = +\infty \quad \text{for} \quad 1 \le i \le n.$$

Together with the trivial inequality

$$0 < a_n = H_0(r) \quad \text{for all} \quad r \in \mathbb{R},$$

it follows easily that there exists a positive real number r with $H_i(r) \ge 0$ for $1 \le i \le n$, that is, $r \in A(P)$.

Now let $r_1, r_2 \in \mathbb{R}$ with $r_1 \in A(P)$ and $r_1 < r_2$. From Lemma 3.8 and the fact that $r_1 \in A(P)$ we may conclude that

$$0 \le H_i(r_1) < H_i(r_2) \quad \text{for} \quad 1 \le i \le n.$$

Together with the trivial inequality $0 < a_n = H_0(r_2)$, it follows that $r_2 \in A(P)$. □

Corollary 3.10 If the leading coefficient of P is positive, then the set $A(P)$ of admissible values for P is either the open interval $(\,0, \infty\,)$, or else it is a half-open interval of the form

$$[\,s, \infty\,) \quad \text{with} \quad 0 < s \in \mathbb{R}.$$

Proof This follows easily from Proposition 3.9 together with the completeness of the real numbers and the fact that all elements of the Horner sequence of P are continuous functions. The fact that the interval is open instead of half-open when the lower bound is zero stems from the fact that admissible values are positive by definition. □

Proposition 3.11 As a function from the real numbers to the real numbers, the polynomial P is strictly increasing on the set $A(P)$ of admissible values for P.

Proof Let $r_1, r_2 \in \mathbb{R}$ with $r_1, r_2 \in A(P)$ and $r_1 < r_2$. By Lemma 3.8, we have $H_i(r_1) < H_i(r_2)$ for $1 \le i \le n$. In particular, $H_n(r_1) < H_n(r_2)$, and thus $P(r_1) < P(r_2)$ by Lemma 3.2. □

Proposition 3.12 Let $r_1, r_2 \in \mathbb{R}$ with $r_1 > 0$, $r_1 \notin A(P)$, and $r_2 \in A(P)$. Then $P(r_1) < P(r_2)$.

Proof We use induction on i to show that

$$H_i(r_1) < H_i(r_2) \quad \text{for} \quad 1 \le i \le n.$$

Then in particular, $H_n(r_1) < H_n(r_2)$, and thus $P(r_1) < P(r_2)$ by Lemma 3.2.

First off, note that $r_1 < r_2$ because the set $A(P)$ of admissible values for P is an upward closed set by Proposition 3.9, and $r_1 \notin A(P)$ and $r_2 \in A(P)$. From one of the assumptions of the proposition, we see that we actually have

$$0 < r_1 < r_2. \tag{10}$$

For the base case of the induction, let $i = 1$. From

$$a_n = H_0(r_1) = H_0(r_2), \quad 0 \leq H_0(r_2), \quad \text{and} \quad a_n \neq 0$$

we obtain $0 < a_n = H_0(r_1) = H_0(r_2)$. We may now use (10) to conclude that

$$H_1(r_1) = H_0(r_1) \cdot r_1 + a_{n-1} < H_0(r_2) \cdot r_2 + a_{n-1} = H_1(r_2).$$

For the induction step, let $i > 1$. Since $r_2 \in A(P)$,

$$0 \leq H_{i-1}(r_2). \tag{11}$$

The induction hypothesis, on the other hand, tells us that

$$H_{i-1}(r_1) < H_{i-1}(r_2). \tag{12}$$

To finish the proof, we need to show, using the relations (10), (11), and (12), that $H_i(r_1) < H_i(r_2)$. We distinguish between two cases.

Case 1: $H_{i-1}(r_1) < 0$
Then, in view of (10) above, $H_{i-1}(r_1) \cdot r_1 < 0$ as well. Furthermore, because of (10) and (11), we have $0 \leq H_{i-1}(r_2) \cdot r_2$, and therefore

$$H_i(r_1) = H_{i-1}(r_1) \cdot r_1 + a_{n-i} < H_{i-1}(r_2) \cdot r_2 + a_{n-i} = H_i(r_2)$$

as desired.

Case 2: $H_{i-1}(r_1) \geq 0$
Then by (12) above, we have $0 \leq H_{i-1}(r_1) < H_{i-1}(r_2)$, and this together with (10) implies that

$$H_i(r_1) = H_{i-1}(r_1) \cdot r_1 + a_{n-i} < H_{i-1}(r_2) \cdot r_2 + a_{n-i} = H_i(r_2). \quad \square$$

We are now ready to give the proof of our main theorem.

Proof of Theorem 3.5 Let r be an admissible value for P that solves Equation (9), and assume for a contradiction that s is a second positive real solution, that is, $P(s) = a$ with $0 < s \in \mathbb{R}$ and $s \neq r$.

From $P(s) = P(r)$ and $s \neq r$ together with the fact that P is strictly increasing on the set $A(P)$ of admissible values for P by Proposition 3.11, we conclude that $s \notin A(P)$. Since s is positive by assumption, Proposition 3.12 now tells us that $P(s) < P(r)$, a contradiction. $\square$

In the next definition and corollary, we investigate what happens when we consider not just positive, but non-negative solutions to Equation (9), and we allow zero as an admissible value.

Definition 3.13 With the notation and assumptions stated preceding Definition 3.3 above, we call the set

$$A_z(P) = \{\, r \in \mathbb{R} \,|\, r \geq 0 \quad \text{and} \quad H_i(r) \geq 0 \quad \text{for} \quad 0 \leq i \leq n \,\}$$

the **set of z-admissible values for *P***. A solution to Equation (9) is called **z-admissible** if it is a z-admissible value for the polynomial P on the left hand side of the equation.

Replacing admissibility with z-admissibility causes no major problems for the theory that we have developed so far. The direct analog to our main result of Theorem 3.5 would be, "z-admissible solutions to Equation (9) are unique as non-negative solutions." That is not quite true, but something only marginally weaker continues to hold.

Corollary 3.14 Suppose that r is a z-admissible solution to Equation (9). Then exactly one of the following holds:

(i) r is the only non-negative solution to Equation (9).

(ii) r is the only positive solution to Equation (9), and zero is a non-z-admissible solution.

Moreover, both (i) and (ii) can actually occur, where in (i), r may be zero or non-zero.

Proof Let r be a z-admissible solution to Equation (9). To see that (i) and (ii) are the only possible cases, suppose that (i) does not hold. Then there is another non-negative solution to Equation (9), say, s, besides r. If r and s were both positive, then they would be two different, admissible solutions to Equation (9), which is impossible by by Theorem 3.5. Therefore, one of r and s must be zero. It remains to show that the z-admissible solution r must be the positive one and the other, that is, zero, cannot be z-admissible. This will easily follow from the following claim: if zero is a z-admissible solution, then there can be no other non-negative solution. To prove the claim, assume that zero is a z-admissible solution. Looking at the definition of z-admissibility, it is easy to see that all non-zero coefficients of P must be positive. It follows that P is strictly increasing on the interval $[\,0, \infty\,)$, and thus there can be no non-negative solution besides zero.

To see that both (i) and (ii) can actually happen, with r being zero or non-zero in (i), let us first recall Example 3.7. Here, $X = 1$ is a solution that is admissible and thus z-admissible, and it is the only non-negative solution. For an example of (i) with $r = 0$, consider the equation

$$X^2 + X = 0.$$

Here, zero and -1 are the only solutions, and zero is z-admissible because

$$\begin{aligned} H_0 &= 1 \\ H_1 &= X+1 \\ H_2 &= X^2+X. \end{aligned}$$

It remains to give an example of (ii) of the theorem. Consider the equation

$$X^2 - X = 0.$$

Here, zero is a solution, but it is not z-admissible since

$$H_1 = X - 1$$

which is negative at zero. The other solution is 1, which is positive and z-admissible because

$$\begin{aligned} H_0 &= 1 \\ H_1 &= X-1 \\ H_2 &= X^2-X. \quad \square \end{aligned}$$

We close this section with some results that are of lesser practical importance but shed some light on Equation (9) and its solutions. We begin with a result that describes the set of right-hand side values for which Equation (9) has an admissible solution.

Proposition 3.15 (Existence of Admissible Solutions) Recall from Corollary 3.10 that if the set of admissible values for P is not empty, then it is either of the form $(\,0,\infty\,)$, or of the form $[\,s,\infty\,)$ with $0 < s \in \mathbb{R}$. In the first case, we have

$$\{\, a \in \mathbb{R} \,|\, P(r) = a \text{ for some } r \in A(P) \,\} = (\,0,\infty\,).$$

In the second case, it is true that

$$\{\, a \in \mathbb{R} \,|\, P(r) = a \text{ for some } r \in A(P) \,\} = [\,P(s),\infty\,),$$

and moreover, $P(s) \geq 0$.

Proof First, note that $P(0) = 0$ since P has a zero constant coefficient. Furthermore, the claim that $P(s) \geq 0$ in the second case follows immediately from the fact that $s \in A(P)$ and $P(s) = H_n(s)$.

Since the set of admissible values for P is not empty, Proposition 3.9 tells us that the leading coefficient of P is positive and therefore,

$$\lim_{r \to +\infty} P(r) = +\infty.$$

Using the intermediate value theorem, the rest of the proposition is now easy to prove from the fact that P is continuous and strictly increasing on $A(P)$. $\square$

Corollary 3.16 If the right-hand side of Equation (9) is negative, then the equation has no admissible solutions. □

The following trivial lemma will allow us to state a slightly stronger version of our main theorem 3.5.

Lemma 3.17 (Existence of Positive Real Solutions) If the leading coefficient of P and the right-hand side of Equation (9) are both positive, then the equation has at least one positive real solution.

Proof Using the intermediate value theorem, this is easy to prove from the fact that $P(0) = 0$ and

$$\lim_{r \to +\infty} P(r) = +\infty. \quad \square$$

Corollary 3.18 (Slightly Stronger Version of Theorem 3.5) There are three possibilities for the positive real solutions to Equation (9):

(i) There is exactly one admissible solution and no other positive solutions,

(ii) there are one or more positive solutions none of which are admissible, or

(iii) there are no positive solutions.

If in addition, we require that the leading coefficient of P and the right-hand side of Equation (9) are both positive, then (iii) is not a possibility.

Proof Let S be the set of positive real solutions to Equation (9). If S is empty, then we're in case (iii), and by Proposition 3.17, this cannot happen when the leading coefficient of P and the right-hand side of the equation are both positive. If S is not empty and does not contain an admissible solution, then we are in case (ii). Else, we are in case (i) by Theorem 3.5.

It remains to show that all situations described in (i), (ii), and (iii) are actually possible. Examples 3.6 and 3.7 cover the case where there is exactly one admissible solution and the one where there are several non-admissible positive solutions. What's missing is an example where there is no positive solution, and one where there is exactly one non-admissible positive solution. Examples with no positive solution are trivially abundant, like $X^2 + 2X = -1$. For an example with exactly one non-admissible positive solution, consider the equation

$$X^3 - 2X^2 + 2X = 1.$$

The only real solution is $X = 1$. This is a non-admissible solution since here, $H_1 = X - 2$, which is negative at $X = 1$. □

Together with Theorem 3.5, the last example of the proof above constitutes a proof of the following corollary:

Corollary 3.19 If r is a positive solution to Equation (9), then r being admissible for P is a sufficient, but not a necessary condition for r to be the only positive solution.

Proposition 3.20 (Monotonicity of Admissible Solutions) Suppose that r_1 is an admissible solution to the equation

$$P(X) = a \quad \text{with} \quad a \in \mathbb{R}$$

and r_2 is an admissible solution to the equation

$$P(X) = b \quad \text{with} \quad b \in \mathbb{R}.$$

Then

$$r_1 < r_2 \iff a < b.$$

Proof This is merely the monotonicity of P on its set of admissible values (Proposition 3.11) reformulated in terms of solutions of equations. □

We close this section with a few remarks on the algorithmic aspects of Theorem 3.5. Given Equation (9), suppose the leading coefficient of P and the right hand side of the equation are both positive. Then there exists a positive solution. One would like to decide whether the unique admissible solution exists, and calculate it if it does.

Naively, we could find an upper bound b for a positive solution, then use binary search on the interval $[\,0, b\,]$ to home in on a positive solution. In the process, we may test, e.g., using Sturm's Theorem, if each Horner polynomial is nonzero on the current search interval. If that is the case, then we can look at the signs of the Horner polynomials on the search interval to obtain a decision on whether an admissible solution exists, and if so, we can refine the binary search to any desired accuracy. But that situation may never occur, namely, when for the true solution r, we have $H_i(r) = 0$ for some Horner polynomial H_i. In that case, the binary search does not lead to a decision. This is one of those situations where approximation to an arbitrary degree of accuracy does not solve the problem at all, because there is a yes-no question involved, and no degree of approximation can provide the answer. Therefore, the only way to decide whether an admissible solution exists is to apply the full force of some real algebraic quantifier elimination algorithm (see e.g. [3] for an overview). Unfortunately, that is not practical at the current time due to the enormous complexity of these algorithms. Luckily, this is not a problem for us at all. For the application that motivated this investigation, it is not Theorem 3.5 for which we need an algorithm. The problem that we must and will solve algorithmically is described in the next section.

4 The Mathematics of the FREQ

In this section, we will prove Theorem 2.3. For the sake of mathematical rigor, the theorem has been rephrased as Theorems 4.8 and 4.10.

As we have mentioned before, the Horner sequence of a polynomial P is used in symbolic computations to efficiently evaluate P. The calculation is based on the recursive definition of the Horner sequence together with the fact that $H_n = P$.

We will now use the same evaluation technique to define a new function that is based on P but acts differently whenever a Horner polynomial becomes negative during the evaluation of P. The idea is to replace the multiplication by X in a Horner polynomial with a multiplication by a constant whenever the previous Horner polynomial evaluates to a negative number. Ultimately, we will be interested in the uniqueness of non-negative real solutions to an equation that is the analogue to Equation (9), except that the polynomial P has been replaced with the modified version that we are about to describe.

Definition 4.1 Let

$$P = \sum_{i=0}^{n} a_i X^i \quad \text{with } a_0, \ldots, a_n \in \mathbb{R} \text{ and } a_n \neq 0,$$

and let $\boldsymbol{b} = (b_1, \ldots, b_n)$ with $b_1, \ldots, b_n \in \mathbb{R}$ and $b_1, \ldots, b_n > 0$.

We now define a sequence

$$K_i : [\,0, \infty\,) \longrightarrow \mathbb{R} \quad (0 \leq i \leq n)$$

of functions recursively as follows:

$$\begin{aligned} K_0(x) &= a_n \\ K_i(x) &= \begin{cases} K_{i-1}(x) \cdot x + a_{n-i} & \text{if } K_{i-1}(x) \geq 0 \\ K_{i-1}(x) \cdot b_i + a_{n-i} & \text{if } K_{i-1}(x) < 0 \end{cases} \qquad \text{for } 1 \leq i \leq n. \end{aligned}$$

We will call the functions defined above the **modified Horner functions for $\boldsymbol{P}$ with respect to $\boldsymbol{b}$**. Of particular interest is the n-th modified Horner function of P. We therefore use the notation

$$\begin{aligned} P_{\boldsymbol{b}}\colon \quad [\,0, \infty\,) &\longrightarrow \mathbb{R} \\ P_{\boldsymbol{b}}(x) &= K_n(x) \end{aligned}$$

and call $P_{\boldsymbol{b}}(x)$ the **negative-constant modification of $\boldsymbol{P}$** with respect to $\boldsymbol{b}$.

Note that the recursive definition of the sequence $K_0, \ldots, K_n$ provides a recipe to calculate any function value $P_{\boldsymbol{b}}(r) = K_n(r)$.

Recall that in the context of fixed-rate investments, the Horner polynomial H_i represents the balance of a fixed-rate investment after i periods. The same is true for the modified Horner functions, except that the modified Horner functions represent a more realistic model of a fixed-rate investment: for each i, the constant b_i is the interest rate charged for borrowing money during the i-th time period, and that rate gets applied whenever the balance is negative.

It is easy to see that in the the case distinctions that occur in the definition above, we could just as easily turn the non-strict inequality in the first case into a strict one, and make the strict inequality in the second case a non-strict one. In the applications that we know of, the version given above is the one that one wants to see. For some of the proofs that follow, however, the other version is more convenient. Therefore, we state it as a lemma.

Lemma 4.2 The definition of the modified Horner functions $K_0, \ldots, K_n$ given in Definition 4.1 above is equivalent to the following definition:

$$\begin{aligned} K_0(x) &= a_n \\ K_i(x) &= \begin{cases} K_{i-1}(x) \cdot x + a_{n-i} & \text{if } K_{i-1}(x) > 0, \\ K_{i-1}(x) \cdot b_i + a_{n-i} & \text{if } K_{i-1}(x) \leq 0. \end{cases} \qquad \text{for } 1 \leq i \leq n. \quad \square \end{aligned}$$

The following lemma shows how the modified Horner functions evaluate to the ordinary Horner sequence when the modified Horner functions are all non-negative.

Lemma 4.3 Let P, $\boldsymbol{b}$, $K_0, \ldots, K_n$, and $P_{\boldsymbol{b}}$ be as in Definition 4.1, and let $r \in [\,0, \infty\,)$. Then, with the notation of Definition 3.1, the following hold:

(i)
$$K_i(r) \geq 0 \quad \text{for } 0 \leq i < n$$
if and only if
$$H_i(r) \geq 0 \quad \text{for } 0 \leq i < n.$$

(ii) If the equivalent conditions of (i) are satisfied, then $K_i(r) = H_i(r)$ for $0 \leq i \leq n$, and $P_{\boldsymbol{b}}(r) = P(r)$.

Proof It is easy to prove by induction on i that each of the two conditions of (i) implies
$$K_i(r) = H_i(r) \quad \text{for } 0 \leq i \leq n.$$

The two conditions of (i) are now easily seen to be equivalent, and the first part of (ii) holds. That together with Lemma 3.2 implies the second part of (ii):

$$P_{\boldsymbol{b}}(r) = K_n(r) = H_n(r) = P(r). \quad \square$$

Next, we prove a sequence of lemmas and propositions that will be instrumental in the proof of our main result, Theorem 4.8.

Lemma 4.4 Let P, $\boldsymbol{b}$, and $K_0, \ldots, K_n$ be as in Definition 4.1. Suppose there exists $0 \leq r \in \mathbb{R}$ and $m \in \mathbb{N}$ with $0 \leq m \leq n$ and

$$K_i(r) \leq 0 \quad \text{for } 0 \leq i < m.$$

Then the following hold:

(i) The premise of the lemma holds for any non-negative real number, that is, $K_i(t) \leq 0$ for all $t \in [\,0, \infty\,)$ and $0 \leq i < m$.

(ii) The function

$$K_m(x) : [\,0, \infty\,) \longrightarrow \mathbb{R}$$

is constant.

Proof It is easy to see that it suffices to prove that the functions

$$K_i(x) : [\,0, \infty\,) \longrightarrow \mathbb{R} \quad (0 \leq i \leq m)$$

are all constant. Let $0 \leq s \in \mathbb{R}$. We will use induction on i to prove that $K_i(s) = K_i(r)$ for $0 \leq i \leq m$. The base case of the induction is trivial because

$$K_0(x) : [\,0, \infty\,) \longrightarrow \mathbb{R}$$

is constant by definition. (The fact the we are not using the premise of the lemma in this case is what makes the proposition work for $m = 0$.) For the induction step, let $i > 0$ and assume that the functions

$$K_j(x) : [\,0, \infty\,) \longrightarrow \mathbb{R} \quad (0 \leq j < i)$$

are all constant. From Lemma 4.2, the induction hypothesis, and the fact that $K_{i-1}(r) \leq 0$ we may conclude that

$$K_i(s) = K_{i-1}(s) \cdot b_i + a_{n-i} = K_{i-1}(r) \cdot b_i + a_{n-i} = K_i(r). \quad \square$$

It is easy to see that in practice, when it comes to checking the condition $K_i(r) \leq 0$ for $0 \leq i < m$ of the lemma above, it suffices to check the seemingly weaker condition

$$K_i(r) \leq 0 \quad \text{for all } i \text{ with } 0 \leq i < m \text{ and } a_{n-i} \neq 0.$$

Proposition 4.5 Let P, $\boldsymbol{b}$, $K_0, \ldots, K_n$, and $P_{\boldsymbol{b}}$ be as in Definition 4.1. If there exists a non-negative real number r with $K_i(r) \leq 0$ for $1 \leq i < n$, then

$$K_i(t) \leq 0 \quad \text{for all } t \in [\,0, \infty\,) \text{ and } 1 \leq i < n,$$

and the function

$$P_{\boldsymbol{b}} : [\,0, \infty\,) \longrightarrow \mathbb{R}$$

is constant.

Proof This is obvious from Lemma 4.4, applied with $m = n$, together with the definition of $P_{\boldsymbol{b}}$, the negative-constant modification of P with respect to $\boldsymbol{b}$. □

Proposition 4.6 Let P, $\boldsymbol{b}$, $K_0, \ldots, K_n$, and $P_{\boldsymbol{b}}$ be as in Definition 4.1. Assume that there exists $0 \leq r \in \mathbb{R}$ and $i \in \mathbb{N}$ with $0 \leq i < n$ such that $K_i(r) > 0$. Then the following hold:

(i) There exists $j \in \mathbb{N}$ with $0 \leq j < n$ and $K_j(t) > 0$ for all $0 \leq t \in \mathbb{R}$.

(ii) $P_{\boldsymbol{b}} : [\,0, \infty\,) \longrightarrow \mathbb{R}$ is strictly increasing.

(iii) $P_{\boldsymbol{b}} : [\,0, \infty\,) \longrightarrow \mathbb{R}$ is unbounded.

Proof Throughout, let m be minimal with the property $K_m(r) > 0$. Then by the assumption of the lemma, we have $0 \leq m < n$.

(i) By Lemma 4.4, the function

$$K_m(x) : [\,0, \infty\,) \longrightarrow \mathbb{R}$$

is constant, and hence $K_m(t) > 0$ for every t with $0 \leq t \in \mathbb{R}$.

(ii) To prove that

$$P_{\boldsymbol{b}} : [\,0, \infty\,) \longrightarrow \mathbb{R}$$

is strictly increasing, assume that $r_1, r_2 \in [\,0, \infty\,)$ with $r_1 < r_2$. We will use induction on i to show that

$$K_i(r_1) < K_i(r_2) \quad \text{for } m + 1 \leq i \leq n.$$

Then in particular,

$$P_{\boldsymbol{b}}(r_1) = K_n(r_1) < K_n(r_2) = P_{\boldsymbol{b}}(r_2).$$

For the base case of the induction, let $i = m + 1$. From our choice of m and Lemma 4.4, we see that $0 < K_m(r) = K_m(r_1) = K_m(r_2)$, and thus

$$\begin{aligned} K_i(r_1) &= K_{i-1}(r_1) \cdot r_1 + a_{n-i} \\ &< K_{i-1}(r_2) \cdot r_2 + a_{n-i} \\ &= K_i(r_2). \end{aligned}$$

For the induction step, let $i > m+1$ and assume that $K_{i-1}(r_1) < K_{i-1}(r_2)$. We distinguish between three cases.

Case 1: $K_{i-1}(r_1) < 0$ and $K_{i-1}(r_2) < 0$
Then

$$\begin{aligned} K_i(r_1) &= K_{i-1}(r_1) \cdot b_i + a_{n-i} \\ &< K_{i-1}(r_2) \cdot b_i + a_{n-i} \\ &= K_i(r_2). \end{aligned}$$

Case 2: $K_{i-1}(r_1) < 0$ and $K_{i-1}(r_2) \geq 0$
Then $K_{i-1}(r_1) \cdot b_i < 0$ and $K_{i-1}(r_2) \cdot r_2 \geq 0$, and thus

$$\begin{aligned} K_i(r_1) &= K_{i-1}(r_1) \cdot b_i + a_{n-i} \\ &< K_{i-1}(r_2) \cdot r_2 + a_{n-i} \\ &= K_i(r_2). \end{aligned}$$

Case 3: $K_{i-1}(r_1) \geq 0$ and $K_{i-1}(r_2) > 0$
Then

$$\begin{aligned} K_i(r_1) &= K_{i-1}(r_1) \cdot r_1 + a_{n-i} \\ &< K_{i-1}(r_2) \cdot r_2 + a_{n-i} \\ &= K_i(r_2). \end{aligned}$$

(iii) To show that $P_{\boldsymbol{b}} : [\,0, \infty\,) \longrightarrow \mathbb{R}$ is unbounded, we will show, by induction on i, that

$$\lim_{x \to +\infty} K_i(x) = +\infty \quad \text{for } m+1 \leq i \leq n.$$

For the base of the induction, let $i = m+1$. By our choice of m and Lemma 4.4, the function

$$K_{i-1}(x) : [\,0, \infty\,) \longrightarrow \mathbb{R}$$

is constant, and its constant value is positive. Therefore,

$$K_i(x) = K_{i-1}(x) \cdot x + a_{n-i},$$

and it follows that

$$\lim_{x \to +\infty} K_i(x) = +\infty.$$

For the induction step, let $i > m+1$ and assume that

$$\lim_{x \to +\infty} K_{i-1}(x) = +\infty.$$

Then there exists a real number t such that $K_{i-1}(s) > 0$ for all $s \in \mathbb{R}$ with $s > t$, and thus

$$K_i(s) = K_{i-1}(s) \cdot s + a_i \quad \text{for all } s \in \mathbb{R} \text{ with } s > t.$$

It is now clear that

$$\lim_{x \to +\infty} K_i(x) = +\infty. \quad \square$$

Lemma 4.7 Let P and $\boldsymbol{b}$ be as in Definition 4.1. Then $P_{\boldsymbol{b}}$, the negative-constant modification of P with respect to $\boldsymbol{b}$, is continuous on the non-negative real numbers.

Proof It is a slightly tedious but elementary exercise to show by induction on i that the modified Horner functions $K_0, \ldots, K_n$ are all continuous on the non-negative real numbers. The essence of the proof lies in the fact that the modified Horner functions are defined by a case distinction, where the distinguishing condition is whether a continuous function is negative or non-negative. Moreover, each branch of the case distinction is a continuous function, and, as noted in Lemma 4.2, the function values agree at each crossover point. In particular, $P_{\boldsymbol{b}} : [\,0, \infty\,) \longrightarrow \mathbb{R}$ is continuous. □

We are now in a position to state and prove and the main theorem of this section. It is an analogue to Theorem 3.5, with the polynomial P in Equation (9) being replaced with the negative-constant modification of P. As in the previous section, we will sometimes use the terms "positive solution" and "non-negative solution" as synonyms for"positive real solution" and "non-negative real solution," respectively. Also, note that while the lemmas and propositions above were proved for arbitrary polynomials with real coefficients, the polynomial P in the theorem below has a zero constant coefficient.

Theorem 4.8 Let

$$P = \sum_{i=0}^{n} a_i X^i \quad \text{with } a_0, \ldots, a_n \in \mathbb{R},\ a_n \neq 0, \text{ and } a_0 = 0.$$

Furthermore, let $\boldsymbol{b} = (b_1, \ldots, b_n)$ with $b_1, \ldots, b_n \in \mathbb{R}$ and $b_1, \ldots, b_n > 0$, let $K_0, \ldots, K_n$ be the modified Horner functions for P, and let $P_{\boldsymbol{b}}(x)$ be the negative-constant modification of P with respect to $\boldsymbol{b}$.

Consider the equation

$$P_{\boldsymbol{b}}(x) = a, \quad \text{where } a \in \mathbb{R}. \tag{13}$$

(i) Assume that

$$K_i(r) > 0 \quad \text{for some } r \in [\,0, \infty\,) \text{ and some } i \text{ with } 0 \leq i < n. \tag{14}$$

Then there exists a $j \in \mathbb{N}$ with $0 \leq j < n$ such that $K_j(t) > 0$ for every $0 \leq t \in \mathbb{R}$, meaning that in practice, it suffices to test the premise (14) on an arbitrarily chosen non-negative real number. Moreover, Equation (13) has exactly one non-negative solution if $a \geq P_{\boldsymbol{b}}(0)$, no non-negative solution otherwise.

(ii) Suppose

$$K_i(r) \leq 0 \quad \text{for some } r \in [\,0, \infty\,) \text{ and all } i \text{ with } 0 \leq i < n. \tag{15}$$

Then $K_i(t) \leq 0$ for all $t \in [\,0, \infty\,)$ and $0 \leq i < n$, meaning that in practice, the premise (15) may be tested on an arbitrarily chosen non-negative real number. Moreover, (13) has a non-negative real solution if and only if $a = P_{\boldsymbol{b}}(0)$. In that case, every non-negative real number is a solution.

(iii) Let $s \in \mathbb{R}$. Then the following are equivalent:

(a) s is an admissible solution (in the sense of Definition 3.3) of the equation $P(x) = a$.

(b) s is a positive solution of Equation (13) and $K_i(s) \geq 0$ for all i with $0 \leq i < n$.

Moreover, if the equivalent conditions (a) and (b) are satisfied, then s is the only non-negative solution to Equation (13).

Proof

(i) follows easily from Proposition 4.6 and Lemma 4.7. (ii) is immediate from Proposition 4.5. The equivalence of conditions (a) and (b) in (iii) is an easy consequence of Lemma 4.3 together with the definition of an admissible solution and the definition of the negative-constant modification of P. If these equivalent conditions are satisfied, then in particular, $K_0(s) \geq 0$. This together with $K_0(s) = a_n \neq 0$ implies that $K_0(s) > 0$. By (i), Equation (13) has at most one non-negative solution, and therefore, s is the only non-negative solution. □

The following corollary provides a sufficient condition for the existence of the unique non-negative solution of Equation (13). The reason for us to state this result is that testing for the condition is trivial by inspection of the equation. Moreover, we have found this to be a common situation in applications of our theory.

Corollary 4.9 Let P, $\boldsymbol{b}$, $P_{\boldsymbol{b}}$, and $K_0, \ldots, K_n$ be as in Theorem 4.8. Assume further that the leading coefficient a_n of P is positive and the right-hand side a of Equation (13) is non-negative. Then Equation (13) has exactly one non-negative real solution.

Proof Since $0 < a_n = K_0(x)$, the premise of Theorem 4.8 (i) is satisfied. It now suffices to show that $a \geq P_{\boldsymbol{b}}(0)$. To this end, we show that $P_{\boldsymbol{b}}(0) \leq 0$.

If $K_{n-1}(0) < 0$, then

$$P_{\boldsymbol{b}}(0) = K_n(0) = K_{n-1}(0) \cdot b_n + a_0 = K_{n-1}(0) \cdot b_n < 0.$$

If $K_{n-1}(0) \geq 0$, then

$$P_{\boldsymbol{b}}(0) = K_n(0) = K_{n-1}(0) \cdot 0 + a_0 = K_{n-1}(0) \cdot 0 = 0. \quad \Box$$

The following theorem provides us with a fully algorithmic version of Theorem 4.8.

Theorem 4.10 There is an algorithm for solving Equation (13) in the following sense: the input to the algorithm is a polynomial P as in Theorem 4.8, a sequence $\boldsymbol{b}$ of n positive real numbers, and a real number a. The output of the algorithm consists of:

(i) An integer variable `state` that has exactly one of the values 0, 1, and 2 such that

$$\begin{aligned}
\texttt{state} = 0 &\iff \text{Equation (13) has a unique non-negative solution}\\
\texttt{state} = 1 &\iff \text{every non-negative number is a solution to Equation (13)}\\
\texttt{state} = 2 &\iff \text{Equation (13) has no non-negative solution}
\end{aligned}$$

(ii) If the value of the variable `state` is 0, the unique non-negative real solution to Equation (13), with arbitrary precision that is limited only by the representation of real numbers that the computing environment uses.

Proof It is immediate from (i) and (ii) of Theorem 3.5 that the following definition of the variable `state` has the desired properties:

$$\texttt{state} = \begin{cases} 0 & \text{if } K_i(0) > 0 \text{ for some } i \text{ with } 1 \leq i < n \text{ and } P_{\boldsymbol{b}}(0) \leq a \\ 1 & \text{if } K_i(0) \leq 0 \text{ for all } i \text{ with } 1 \leq i < n \text{ and } P_{\boldsymbol{b}}(0) = a \\ 2 & \text{otherwise} \end{cases}$$

If the value of `state` is 0, then by Proposition 4.6 and Lemma 4.7, $P_{\boldsymbol{b}}$ is strictly increasing, continuous, and unbounded on the interval $[\,0, \infty\,)$. Therefore, approximating the unique non-negative solution of Equation (13) with arbitrary precision is a simple exercise of finding an upper bound and then performing a binary search. □

A few remarks on the practicalities of implementing the algorithm of Theorem 4.10 are in order. In a sense, the algorithm combines the best of the two worlds of symbolic and numeric computation. Unlike most symbolic algorithms, this one does not require the use of unlimited precision arithmetic. It can be implemented using floating point arithmetic, as long as one is willing to accept the limited precision, and the possibility of numerical overflow or underflow in the occasional corner case. There is no reason to believe that these pathological cases will be any worse or more frequent here than they are in any floating point calculation. In this sense, the calculation is numerical in nature. On the other hand, the algorithm does not come with the kind of uncertainties that numerical methods sometimes have, like Newton's method failing to converge. If one does indeed use unlimited precision arithmetic (often referred to as BigInt and BigRational in programming libraries), then our criterion for the existence of a unique non-negative solution is unambiguous, and if the criterion is met, the binary search will approximate the solution with arbitrary precision. In this sense, the algorithm is symbolic in nature.

It is perhaps noteworthy that none of the proofs in this section used any results from the previous section. In particular, our proof of Theorem 4.8 is entirely independent of the proof of Theorem 3.5 that was given in the previous section. We now show how Theorem 3.5 can be proved from Theorem 4.8. We believe, however, that this is just a curiosity; for a full understanding of our topic, it is well worth having all the results of the previous section.

So for an alternate proof of Theorem 3.5, let r be an admissible solution to the equation

$$P(x) = a,$$

and assume for a contradiction that s is another positive solution to the equation. By (iii) of Theorem 4.8, r is the unique non-negative solution to Equation (14) for any sequence $\boldsymbol{b}$ of positive real numbers. In particular, this is true for the sequence $\boldsymbol{s}$ whose elements all equal s. But it is easy to see from Definition 4.1 that

$$P_{\boldsymbol{s}}(s) = P(s),$$

and therefore, s is a second non-negative solution to Equation (14), a contradiction.

References

[1] Daniel Teichroew, Alexander A. Robicheck, and Michael Montalbano, *Mathematical analysis of rates of return under certainty.* Management Science Vol. 11 Nr. 3, January 1965, 395–403.

[2] Daniel Teichroew, Alexander A. Robicheck, and Michael Montalbano, *An analysis of criteria for investment and financing decisions under certainty.* Management Science Vol. 12 Nr. 3, November 1965, 151–179.

[3] Andreas Dolzmann, Thomas Sturm, and Volker Weispfenning, *Real Quantifier Elimination in Practice*, Technical Report MIP-9720, University of Passau, Germany, 1997. Online at andreas.dolzmann.de/publications/aant.pdf

[4] http://en.wikipedia.org/wiki/Internal_rate_of_return#Calculation