Abstract CRSP+ Databases

Incomplete.

Let D be a CRSP+ database.

Pricing

For any pair of integers P, d $\in \mathbb{Z}$, we will say that **D prices P on d** if and only if the set {|R.prc| | R is a record in D's Price table where R.permno = P and R.date = d} contains exactly one non-zero numerical value, which we call **D's price for P on d**. Note that by this definition, D may price P on d even if there is more than one record, R, in D's prices table where R.permno = P and R.date = d, so long as P's price itself on d is unambiguous when numerical signs are ignored.

If S $\subset \mathbb{Z}$, we will say that **D prices P in S** if and only if there exists at least one d $\in$ S such that D prices P on d. Likewise, we will say that **D does not price P in S** if and only if there exists no d $\in$ S such that D prices P on d.

Market Dates and Market Intervals

Define $\mathbb{M}$, **the set of market dates in D**, to be the set of all integers 0 < m < 99991231 for which there exists at least 100 integers, P, such that D prices P on m. Define **$\mathbb{M}$***, **the extended market dates in D**, to be the set $\mathbb{M} \cup \{0, 99991231\}$. Order $\mathbb{M}$* by the ordering it inherits as a subset of the integers. Since $\mathbb{M}$* is a finite ordered set, the set-functions max and min are well-defined on the nonempty subsets of $\mathbb{M}$*.

If a and b are extended market dates in D, define

> **[a, b]** = $\{m \in \mathbb{M}^* \mid a \le m \le b\}$
> **[a, b)** = $\{m \in \mathbb{M}^* \mid a \le m < b\}$
> **(a, b]** = $\{m \in \mathbb{M}^* \mid a < m \le b\}$
> **(a, b)** = $\{m \in \mathbb{M}^* \mid a < m < b\}$

If the subset I $\subset \mathbb{M}$* can be expressed in one of these four forms (and the form will usually *not* be unique), then I is said to be an **extended market interval in D**. An extended market interval I is defined to be a **market interval in D** if and only if I $\subset \mathbb{M}$.

Note that extended market intervals of type (2) and (4) cannot include 99991231, while those of type (3) and (4) cannot include 0. Those of type (4) therefore contain neither 99991231 nor 0 and are thus always market intervals in D, even if a or b is in {0, 99991231}.

For each m $\in \mathbb{M}$* define **m + 1** = min((m, 99991231] $\cup$ {99991231}) and **m – 1** = max([0, m) $\cup$ {0}). Note that by this definition, 99991231 + 1 = 99991231 and 0 – 1 = 0. These "increment" and "decrement" operations thus defined are consistent with the ordering on $\mathbb{M}$* in the sense that the following properties, which should be obvious, hold:

1. If x < 99991231 then x < x + 1.
2. If x > 0 then x > x – 1.

3. If $y > x$ then $y \geq x + 1$.
4. If $y < x$ then $y \leq x - 1$.

The following lemma merely states a property of any finite ordered set. While versions of this lemma are true with any of the four interval representations listed above, we state and prove only one of them but use the others without mention:

Lemma 1: If $[a, b)$ and $[c, d)$ are extended market intervals in D, then $[a, b) \cap [c, d) = [\max(a, c), \min(b, d))$, which is empty if $\max(a, c) \geq \min(b, d)$.

> *Proof*: $[a, b) \cap [c, d) = \{x \in \mathbb{M}^* \mid a \leq x < b \text{ and } c \leq x < d\} = \{x \in \mathbb{M}^* \mid a \leq x \text{ and } c \leq x\} \cap \{x \in \mathbb{M}^* \mid x < b \text{ and } x < d\} = \{x \in \mathbb{M}^* \mid \max(a, c) \leq x\} \cap \{x \in \mathbb{M}^* \mid x < \min(b, d)\} = \{x \in \mathbb{M}^* \mid \max(a, c) \leq x \text{ and } x < \min(b, d)\} = [\max(a, c), \min(b, d))$.

"Cash" Data

If D prices the integer 0 in $\mathbb{M}$, then for each $d \in \mathbb{M}$ where D does not price 0 on d,

1. Set $d' = \max(\{m \in \mathbb{M} \mid m < d \text{ and D prices 0 on } m\} \cup \{\min(\{m \in \mathbb{M} \mid \text{D prices 0 on } m\})\})$.
2. Delete all records, R, in D's prices table with R.permno = 0 and R.date = d.
3. Add exactly one record, R, to D's prices table with R.permno = 0, R.date = d, and R.prc = p, where p is D's price for 0 on d'.

If D does not price 0 in $\mathbb{M}$, then for each $d \in \mathbb{M}$, add exactly one record, R, to D's prices table with R.permno = 0, R.date = d, and R.prc = 1.

In either case, we have ensured that D prices the integer 0 (used to store performance history on cash investments) on every market date in D. Finally, add exactly one record, R, to D's Header table, setting R.permno = 0 and R.permco = 0.

Permno's and Permco's

Define $\mathcal{P}$, **the set of permno's in D**, to be the set of all integers, P, appearing in the field permno of D's Header table and having the property that D prices P on d for some market date $d \in \mathbb{M}$. Define $\mathcal{C}$**, the set of permco's in D**, to be the set of all integers, C, such that there exists a record, R, in D's Header table with R.permco = C and R.permno $\in \mathcal{P}$. Note that the last section guarantees that $0 \in \mathcal{P}$ and $0 \in \mathcal{C}$. The following axiom concerning D's Header table must be empirically verified:

Axiom 1: For each $P \in \mathcal{P}$ there exists exactly one record R in D's Header table with R.permno = P.

Define the function **permco**: $\mathcal{P} \mapsto \mathcal{C}$, **D's permco function**, as follows: For each $P \in \mathcal{P}$, let permco(P) = R.permco, where R is the record (unique by Axiom 1) where R.permno = P.

The π Mapping

Define $\boldsymbol{\pi}$**:** $\mathcal{P} \times \mathbb{M}^* \mapsto \mathbb{M}^*$ (where π is for "previous") by

$$\pi(P, d) = \max(\{x \in \mathbb{M}^* \mid x \leq d \text{ and D prices P on } x\} \cup \{0\})$$

for each $(P, d) \in \mathcal{P} \times \mathbb{M}^*$. In other words, given an extended market date d, $\pi(P, d - 1)$ is the "previous" extended market date on which D prices P, if one exists, or else 0.

The following proposition lists some useful properties of the π mapping:

Proposition 2: For all $(P, d) \in \mathcal{P} \times \mathbb{M}^*$ and $x, y \in \mathbb{M}^*$:

1. Either D prices P on $\pi(P, x)$ or $\pi(P, x) = 0$.
2. $x \geq \pi(P, x) \geq \pi(P, x - 1)$.
3. $x = \pi(P, x)$ if and only if $x = 0$ or D prices P on x.
4. $x = \pi(P, x - 1)$ if and only if $x = 0$.
5. If $y \leq x$ and D prices P on y then $y \leq \pi(P, x)$.
6. D does not price P in $(\pi(P, x - 1), x)$.

Proof: For each $d \in \mathbb{M}^*$ define $A_d = \{m \in \mathbb{M}^* \mid m \leq d \text{ and D prices P on } m\} \cup \{0\}$. By definition, then, $\pi(P, d) = \max(A_d)$. Fix any $x \in \mathbb{M}^*$.

Property (1) follows because A_x, being a finite set, always contains its maximum, $\pi(P, x)$. The first inequality in (2) follows because $x \geq y$ for all $y \in A_x$ and $\pi(P, x) \in A_x$. The second half of (2) follows because $A_x \supset A_{x-1}$ and because the max function decreases as its argument decreases. (5) merely asserts that if $y \in A_x$ then $\max(A_x) \geq y$, which is trivial.

Note that (3) simply asserts that $x = \pi(P, x)$ if and only if $x \in A_x$. If $x = \pi(P, x)$ then $x \in A_x$ follows from (1). Conversely if $x \in A_x$ then $\pi(P, x) = \max(A_x) \geq x$, which, combined with (2) implies $\pi(P, x) = x$.

For (4), if $x = 0$ then $x = \pi(P, x - 1)$ follows from (2). Assume $x = \pi(P, x - 1)$, that is, $x = \max(A_{x-1})$. This implies $x \in A_{x-1}$. But $x \leq x - 1$ only if $x = 0$.

If $x > y$ and D prices P on y, then (5) implies $\pi(P, x - 1) \geq y$. Hence (6).

Reconciling the Event Tables with $\mathbb{M}^*$

Before proceeding, some modifications to D may be necessary to make its event data consistent with $\mathbb{M}^*$, our (extended) calendar of valid market dates deduced from D's Price table. To do so, we will need the following funciton: Define

$\boldsymbol{\upsilon}$**:** $\mathbb{Z} \mapsto \mathbb{M}^*$ by $\upsilon(x) = \min(\{m \in \mathbb{M}^* \mid x \leq m\} \cup \{99991231\})$ for all $x \in \mathbb{Z}$.

In words, υ takes any integer x and moves it "up" (in the quantitative sense, not in the colloquial sense of moving dates "up" on the calendar) to the next extended market date, if it isn't one already, and returns x otherwise.

Step 1: Within each of the five fields in D listed below, replace any values that are not integers greater than zero with the value 99991231:

[Distributions].exdt
[Delistings].dlstdt
[Delistings].nextdt
[Names].namedt
[Shares].shrsdt

Step 2: Within each of the four fields listed below, replace any field value d (which is a positive integer as of Step 1) where $d \notin \mathbb{M}^*$ with the value $\upsilon(d)$:

[Distributions].exdt
[Delistings].nextdt
[Names].namedt
[Shares].shrsdt

Step 3: Add one record, R, to D's Delistings table, setting

R.permno = 0
R.dlstdt = max($\mathbb{M}$)
R.dlstcd = 100
R.nwperm = R.nwcomp = 0
R.nextdt = 99991231
R.dlprc = 0
R.dlret = –88

Step 4: For each record, R, in D's Delistings table with R.dlstdt = max($\mathbb{M}$), reset R.dlstdt = π(R.permno, max($\mathbb{M}$)).

Step 5: For each distinct $(P, d) \in \mathcal{P} \times \mathbb{M}$ such that

1. there exists a record, R, in D's Names table with P = R.permno, d = π(R.permno, R.namedt – 1), and R.exchcd = 0, but
2. there exists no record, R, in D's Delisting table with R.permno = P and R.dlstdt = d,

add exactly one record, S, to D's Delistings table, setting

S.permno = P
S.dlstdt = d
S.dlstcd = 800
S.nwperm = S.nwcomp = 0
S.nextdt = 99991231
S.dlprc = 0
S.dlret = –55

Delisting Dates and Trailer Periods

Having made the modifications in the last section, a market date d is defined to be a **delisting date** for P $\in \mathcal{P}$ if there exists a record, R, in D's Delistings table with R.permno = P and R.dlstdt = d. Thus by Step 3 of the previous section, for example, max($\mathbb{M}$) is a delisting date for 0. The following four axioms regarding delisting dates, which should now hold for any CRSP+ database, must be empirically verified:

Axiom 2: max($\mathbb{M}$) is 0's only delisting date.

Axiom 3: If d is a delisting date for P, then D prices P on d but not on d + 1.

Axiom 4: If D prices P on x then there exists a delisting date d for P such that $x \le d$.

Axiom 5: The pair of fields (permno, dlstdt) forms a key for D's Delistings table.

An extended market interval (d, b] is a **trailer period** for P if

1. d is a delisting date for P,
2. D does not price P in (d, b], and
3. either D prices P on b + 1 or b = 99991231.

Note that by Axiom 2, (max($\mathbb{M}$), 99991231] = {99991231} is 0's only trailer period.

Proposition 3: Distinct trailer periods for P are disjoint.

> *Proof*: Suppose (d, b] and (d', b'] are two trailer periods for P containing x. Then since D prices P on both d < x and d' < x, Proposition 2.5 implies $d \le \pi(P, x) \le x$ and $d' \le \pi(P, x) \le x$. But since D prices P in neither (d, x] nor (d', x], it follows that $d = \pi(P, x) = d'$.
>
> If b < b' then $b \neq 99991231$, meaning D prices P on b + 1. But $b + 1 \le b'$, which implies, since d = d', that $b + 1 \in$ (d', b'], a contradiction. Thus $b \ge b'$. The same argument, only with (d, b'] and (d, b] reversed, proves $b' \ge b$. Hence b = b'. We have thus shown that intersecting trailer periods for P are identical, which completes the proof.

Let $\mathbf{L_P}$ be the set of delisting dates for P and let $\mathbf{T_P}$ be the set of trailer periods for P. There is a natural mapping $\boldsymbol{\lambda_P}$**:** $\mathbf{T_P} \mapsto \mathbf{L_P}$ (abbreviated λ: T $\mapsto$ L when the context of a fixed P $\in \mathcal{P}$ is clear) given by λ((d, b]) = d for each (d, b] $\in$ T.

Proposition 4: The mapping λ from T, the set of trailer periods for P, to L, the set of delisting dates for P, given by λ((d, b]) = d for all (d, b] $\in$ T, is a bijection.

> *Proof*: Suppose λ((d, b]) = a = λ((d', b']), that is, d = a = d'. Then both (d, b] and (d', b'] contain a + 1, so by Proposition 3, (d, b] = (d', b']. Thus λ is injective.
>
> To prove that λ is surjective, let d be any delisting date for P and let E = {x $\in \mathbb{M}$ | x > d and D prices P on x}.

If $E \neq \varnothing$, let $b = \min(E) - 1$. Then D prices P on $b + 1$. By Axiom 4, D does not price P on $d + 1$, so $d + 1 < \min(E) = b + 1$, hence $d < b$. Clearly D does not price P in $(d, b]$, so it follows that $(d, b]$ is a trailer period for P and that $\lambda((d, b]) = d$.

If, on the other hand, $E = \varnothing$, let $b = 99991231$. Then D does not price P in $(d, b]$, and thus $(d, b]$ is a trailer period for P such that $\lambda((d, b]) = d$.

Ending Dates

A market date d is defined to be an **ex-dividend date** for $P \in \mathcal{P}$ if there exists a record in D's Distributions table with permno = P, exdt = d, divamt > 0, and distcd ≥ 1000. Note that by definition, ex-dividend dates cannot be 0 or 99991231.

Add a new field to D's Delistings table called **enddt** (for "ending date") and set its values as follows:

For each record, S, in D's Delistings table, make the following modifications to D's event tables: Let P = S.permno, d = S.dlstdt, and $(d, b] = \lambda_P^{-1}(d)$, where the mapping $\lambda_P: T_P \mapsto L_P$ is as defined in Proposition 4. Let p be D's price for P on d, which is guaranteed to exist by Axiom 3.

If S.nextdt ∈ (d, b], S.nextdt < 99991231, and S.dlprc is a numerical value, then

1. Set S.enddt = nextdt.
2. Add exactly one new record, R, to D's Distributions table, setting

   R.permno = P
   R.distcd = S.dlstcd
   R.divamt = |S.dlprc|
   R.exdt = S.enddt

   and all remaining fields in R equal to zero.

Otherwise, let $X = \{x \in (d, b] \mid x \text{ is an ex-dividend date for } P\}$.

If $X = \varnothing$ then

1. Set S.enddt = d + 1 (which could be 99991231).
2. Add add exactly one new record, R, to D's Distributions table, setting

   R.permno = P
   R.distcd = −1 · S.dlstcd

$$\text{R.divamt} = p \cdot \begin{cases} 0 & \text{if S.dlret} = -1 \\ 0.7 & \text{if S.dlret} \neq -1 \text{ and either } \\ & 500 \leq \text{S.dlstcd} < 600 \text{ or S.dlstcd} = 800 \\ 1 & \text{otherwise.} \end{cases}$$

   R.exdt = d + 1

   and all remaining fields in R equal to zero.

Otherwise, set S.enddt = max(X).

Recall that valid distribution codes (as defined by CRSP) are four-digit positive integers, whereas valid delisting codes are three-digit positive integers. Thus, the addition of records to the Distributions table in the above procedure is easily reversible. For the same reasons, the addition of these records has no affect on the identification of ex-dividend dates as defined at the beginning of this section.

After populating the enddt field of D's Delistings table, we define, for each fixed $P \in \mathcal{P}$, a mapping $\boldsymbol{\varepsilon_P}$**:** $\mathbf{L_P} \mapsto \mathbb{M}$***** as follows: For each $d \in L_P$, Axiom 5 assures the existence of exactly one record, R, in D's Delistings table with R.permno = P and R.dlstdt = d. We may therefore define $\varepsilon_P(d)$ = R.enddt.

An extended market date e is defined to be an **ending date** for P if and only if $e = \varepsilon_P(d)$ for some delisting date d for P, or equivalently, if and only if there exists a record in D's Delistings table with permno = P and enddt = e. Let $\mathbf{E_P}$ denote the set of ending dates for P, that is, the range of the mapping ε_P. As with the mapping $\lambda_P: T_P \mapsto L_P$, which we abbreviate with $\lambda: T \mapsto L$ when the context of a fixed $P \in \mathcal{P}$ is clear, we will abbreviate the mapping $\varepsilon_P: L_P \mapsto E_P$ as $\varepsilon: L \mapsto E$ in the same context.

Note that by the procedure above, $\varepsilon(d)$ always falls within the trailer period $\lambda^{-1}(d)$ for any $d \in L$. Since the mapping λ^{-1} is injective (Proposition 4), and since distinct trailer periods for P are disjoint (Proposition 3), this implies that ε is injective. The ending dates for P are by definition the range of ε, so it follows that ε is surjective as a mapping to P's ending dates. Thus $\varepsilon: L \mapsto E$ is a bijection. Finally, note that since $(\max(\mathbb{M}), 99991231] = \{99991231\}$ is 0's only trailer period, 99991231 is its only ending date.

Proposition 5: Fix $P \in \mathcal{P}$ and let T be the set of P's trailer periods, L the set of P's delisting dates, and E the set of P's ending dates. There exist unique mappings

1. $\lambda: T \mapsto L$ such that $\lambda(\ (d, b]\ ) = d$ for each $(d, b]$ in T.
2. $\varepsilon: L \mapsto E$ such that, for each $d \in L$, $\varepsilon(d)$ = R.enddt where R is the unique record in D's Delistings table where R.permno = P and R.dlstdt = d.
3. $\boldsymbol{\tau}: \mathbf{E} \mapsto \mathbf{T}$ such that $e \in \tau(e)$ for each $e \in E$.

Furthermore, these mappings are all bijections.

> *Proof*: Simply let λ be the bijection defined in Proposition 4 and let ε be the mapping used to define P's ending dates, which was shown to be a bijection in the remarks following its definition. λ and ε then satisfy (1) and (2), respectively, by definition. We may then define $\tau: E \mapsto T$ by $\tau = (\varepsilon \circ \lambda)^{-1}$, which is a bijection. If $e \in E$ then $\tau(e) = (\varepsilon \circ \lambda)^{-1}(e) = \lambda^{-1}(\varepsilon^{-1}(e))$ contains e because, as was noted above, $\lambda^{-1}(d)$ always contains $\varepsilon(d)$ for all $d \in L$. Thus τ satisfies (3).
>
> Suppose $\lambda': T \mapsto L$, $\varepsilon': L \mapsto E$, and $\tau': E \mapsto T$ are mappings satisfying (1) through (3), respectively. Then (1) and (2) make it clear that $\lambda' = \lambda$ and $\varepsilon' = \varepsilon$. Since (3) implies that $e \in \tau(e)$ and $e \in \tau'(e)$ for each $e \in E$, Proposition 3 implies that $\tau(e) = \tau'(e)$ for all $e \in E$, and thus that $\tau = \tau'$. This completes the proof.

To summarize, specifying one of either a trailer period, a delisting date or an ending date unambiguously specifies all three in a natural way. The next corollary lists some useful properties of the mappings defined in Proposition 5:

Corollary 6: For a fixed $P \in \mathcal{P}$, with $\lambda: T \mapsto L$, $\varepsilon: L \mapsto E$, and $\tau: E \mapsto T$ as in Proposition 5:

1. Each trailer period for P contains only one ending date for P.
2. D does not price P on its ending dates.

If $d \in L$, $e \in E$, and $d < e$, then

3. $d < \varepsilon(d) \le e$.
4. $(d, \varepsilon(d)] \subset \lambda^{-1}(d)$
5. $d \le \varepsilon^{-1}(e) < e$
6. $(\varepsilon^{-1}(e), e] \subset \tau(e)$
7. $\pi(P, x) = d$ for all $x \in \lambda^{-1}(d)$
8. If $99991231 \in E$ then $\varepsilon^{-1}(99991231) = \max(\mathbb{M})$.

Proof: Suppose $(d, b] \in T$ and $e \in E \cap (d, b]$. Then $e \in \tau(e)$. Since intersecting trailer periods for P are identical, it follows that $\tau(e) = (d, b]$, and hence that $e = \tau^{-1}((d, b])$. Thus $\tau^{-1}((d, b])$ is the only ending date for P contained in (d, b], which proves (1).

If $e \in E$ then $e \in \tau(e)$. Since D does not price P in its trailer periods, (2) follows.

Suppose $d \in L$, $e \in E$, and $d < e$. Let $(d, b] = \lambda^{-1}(d)$. Since $\varepsilon(d)$ falls within $\lambda^{-1}(d) = (d, b]$, (4) immediately follows, as does the fact that $d < \varepsilon(d)$ for all $d \in L$. By (1), $\varepsilon(d)$ is the only ending date for P contained in $(d, b] \supset (d, \varepsilon(d)]$, so (3) follows.

(5) and (6) now follow from (3) and (4): Let $d' = \varepsilon^{-1}(e)$ and $e' = \varepsilon(d)$. By (3), $d' < \varepsilon(d')$, implying $\varepsilon^{-1}(e) < e$. (4) implies that $(\varepsilon^{-1}(e), e] = (d', \varepsilon(d')] \subset \lambda^{-1}(d') = (\varepsilon \circ \lambda)^{-1}(e) = \tau(e)$, where we have used the definition $\tau = (\varepsilon \circ \lambda)^{-1}$ from the proof of Proposition 5. Since $\tau(e) \supset (\varepsilon^{-1}(e), e]$ contains no delisting dates for P, the rest of (5) follows.

Suppose $x \in \lambda^{-1}(d)$. Since D prices P on d and $d < x$, Proposition 2.5 implies $d \le \pi(P, x)$. Thus by Proposition 2.1 and 2.2, D prices P on $\pi(P, x) \le x$. But since D does not price P in $(d, x] \subset \lambda^{-1}(d)$ we conclude $\pi(P, x) = d$, which proves (7).

To prove (8), assume $99991231 \in E$ and let $d = \varepsilon^{-1}(99991231)$. Let R be the record in D's Delisting table with R.permno = P and R.enddt = 99991231. If R.nextdt $\in$ (d, 99991231), then by definition we would have $\varepsilon(d)$ = R.nextdt, a contradiction. Likewise, if (d, 99991231) contained ex-dividend dates for P, then $\varepsilon(d)$ would equal their maximum, also a contradiction. We must therefore conclude that $99991231 = \varepsilon(d) = d + 1$, and hence that $d = \max(\mathbb{M})$.

Starting Dates

Fix any $P \in \mathcal{P}$. Let L be the set of P's delisting dates, E the set of P's ending dates, and T the set of P's trailer periods. Let λ, ε, and τ be defined according to Proposition 5.

Define $\boldsymbol{\nu}\mathbf{:}\ \mathcal{P} \times \mathbb{M}^* \mapsto \mathbb{M}^*$ (where ν is for "next") by

$$\nu(P, d) = \min(\{x \in \mathbb{M}^* \mid x \ge d \text{ and either D prices P on } x,\ x \in E, \text{ or } x = 99991231\})$$

for each $(P, d) \in \mathcal{P} \times \mathbb{M}^*$. In other words, given an extended market date d, $\nu(P, d + 1)$ is the "next" extended market date that is either an ending date for P, a market date on which D prices P, or simply 99991231—whichever comes first.

The next proposition lists some properties of ν that will be useful later:

Proposition 7: For any $P \in \mathcal{P}$ and $x, y \in \mathbb{M}^*$,

1. Either D prices P on $\nu(P, x)$, $\nu(P, x) \in E$, or $\nu(P, x) = 99991231$
2. $x \le \nu(P, x) \le \nu(P, x + 1)$
3. $x = \nu(P, x)$ if and only if either D prices P on x, $x \in E$, or $x = 99991231$
4. $x = \nu(P, x + 1)$ if and only if $x = 99991231$
5. If $x \le y$ and either D prices P on y, $y \in E$, or $y = 99991231$, then $\nu(P, x) \le y$.
6. If $x \le y \le \nu(P, x)$ then $\nu(P, x) = \nu(P, y)$.
7. D does not price P in $(x, \nu(P, x + 1))$.
8. If $x < 99991231$ then $(x, \nu(P, x + 1)) \cap E = \varnothing$.
9. If $d \in L$ then $\nu(P, d + 1) = \varepsilon(d)$.
10. If $x = 0$ or $x \in E$, then either D prices P on $\nu(P, x + 1)$ or $\nu(P, x + 1) = 99991231$.
11. $\nu(P, \nu(P, x)) = \nu(P, x)$.

Proof: For each $d \in \mathbb{M}^*$ define $A_d = \{m \in \mathbb{M}^* \mid m \ge d$ and either D prices P on m, $m \in E$, or $m = 99991231\}$. By definition, then, $\nu(P, d) = \min(A_d)$. Fix any $x \in \mathbb{M}^*$.

Property (1) follows because A_x, being a finite set, always contains its minimum, $\nu(P, x)$. The first inequality in (2) follows because $x \le y$ for all $y \in A_x$ and $\nu(P, x) \in A_x$. The second half of (2) follows because $A_{x+1} \subset A_x$ and the min function decreases as its argument increases. (5) merely asserts that if $y \in A_x$ then $\min(A_x) \le y$, which is trivial.

Note that (3) simply asserts that $x = \nu(P, x)$ if and only if $x \in A_x$. If $x = \nu(P, x)$ then $x \in A_x$ follows from (1). Conversely if $x \in A_x$ then $\nu(P, x) = \min(A_x) \le x$, which, combined with (2) implies $\nu(P, x) = x$.

For (4), if $x = 99991231$ then $x = \nu(P, x + 1)$ follows from (2). Assume $x = \nu(P, x + 1)$, that is, $x = \min(A_{x+1})$. This implies $x \in A_{x+1}$. But $x \ge x + 1$ only if $x = 99991231$.

If $x \le y \le \nu(P, x)$ then (1) applied to $\nu(P, x)$ and (5) applied to $y \le \nu(P, x)$ imply $\nu(P, y) \le \nu(P, x)$. (2) implies $x \le y \le \nu(P, y)$. Thus (1) applied to $\nu(P, y)$ and (5) applied to $x \le \nu(P, y)$ imply $\nu(P, x) \le \nu(P, y)$. Hence (6).

If $x < y$ and D prices P on y, then (5) implies $\nu(P, x + 1) \le y$. Hence (7).

If $x < 99991231$ and $x < e \in E$, then (5) implies $\nu(P, x + 1) \le e$. Hence (8).

For (9), note that if $d \in L$ then $d < \varepsilon(d) \in E$ by Corollary 6.3. Thus (5) implies $\nu(P, d + 1) \le \varepsilon(d)$. (2) and (4) imply (since $d < 99991231$) $d < \nu(P, d + 1)$. Hence $\nu(P, d + 1) \in (d, \varepsilon(d)] \subset \lambda^{-1}(d)$ by Corollary 6.4. Corollary 6.1, the fact that D does not price P in trailer periods, and (1) now imply $\nu(P, d + 1) = \varepsilon(d)$.

To prove (10), assume, to obtain a contradiction, that D does not price P on $\nu(P, x + 1) < 99991231$. By (1), it follows that $\nu(P, x + 1) = e \in E$. By (4), $x < e$. By (7), D does not price P in (x, e), and by Corollary 6.5, $d = \varepsilon^{-1}(e) < e$. Since D prices P on d, we must have $d \le x$. We have shown that $x \in [d, e)$, implying that $x \ne 0$ and, by Corollary 6.1 and 6.4, that $x \notin E$, which is the contradiction we sought.

Property (11) follows from (1) and (3).

We are finally ready for the main definition of this section: A market date s is defined to be a **starting date** for P if $s = \nu(P, e + 1)$, where $e \in \{0\} \cup E$. Note the requirement that s be a market date, which rules 99991231 out from being a starting date for P, and also implies, by Proposition 7.4, that $e < s$. Also note that by Proposition 7.10, D always prices P on its starting dates, as we would like. Thus ending dates for P cannot be starting dates for P.

Note that since D prices 0 on every market date in D and since 99991231 is 0's only ending date, $\nu(0, 0 + 1) = \nu(0, \min(\mathbb{M})) = \min(\mathbb{M})$ is 0's only starting date.

Tracking Periods

Fix any $P \in \mathcal{P}$. Let L be the set of P's delisting dates, E the set of P's ending dates, and T the set of P's trailer periods. Let λ, ε, and τ be defined according to Proposition 5.

An extended market interval [s, e) is defined to be a **tracking period** for P if

1. s is a starting date for P,
2. e is an ending date for P, and
3. (s, e) contains no ending dates for P.

Note that $\mathbb{M}$ is 0's only tracking period.

Proposition 8: If s is a starting date for P, then $s - 1$ does not belong to a tracking period for P.

Proof: By definition, $s = \nu(P, x + 1)$ for some $x \in \{0\} \cup E$. Suppose, to obtain a contradiction, that $s - 1$ belongs to a tracking period [s', e') for P. By Proposition 7.7, D does not price P in (x, s). Since D prices P on s' and $s' < s$, it follows that $s' \leq x$. But since $x \leq s - 1 < e'$, we have $x \in [s', e')$, which is the desired contradiction, since $x \in \{0\} \cup E$.

Corollary 9: If [s, e) is a tracking period for P, then (s, e) contains no starting dates for P.

Proof: If (s, e) contained a starting date s' for P, then [s, e) would contain $s' - 1$, contradicting Proposition 8.

Proposition 10: Distinct tracking periods for P are disjoint.

Proof: Suppose [s, e) and [s', e') are tracking periods for P with nonempty intersection. Then by Lemma 1, $g = \max(s, s')$ is a starting date for P contained in both [s, e) and [s', e'). Since (s, e) contains no starting dates for P, we must have $g = s$. But (s', e') contains no starting dates for P, so $g = s'$. Hence $s = s'$.

Similarly, $h = \min(e, e')$ is an ending date for P contained in $[s, e] \cap [s', e']$. Since $[s, e)$ contains no ending dates for P, $h = e$, and since $[s', e')$ contains no ending dates for P, $h = e'$. Thus $e = e'$.

Proposition 11: If D prices P on x then x belongs to a tracking period for P (which is unique by the last proposition).

Proof: By Axiom 4, there exists some $d \in L$ such that $x \le d$. Thus $B = \{y \in \mathbb{M}^* \mid x < y \text{ and } y \in E\}$ contains $\varepsilon(d) > d$ and is thus nonempty, so let $e = \min(B)$. Clearly $x < e$, $e \in E$, and $[x, e) \cap E = \varnothing$ ($x \notin E$ because D prices P on x).

Let $A = ([0, x) \cap E) \cup \{0\}$. The set A is nonempty, so let $a = \max(A)$. Then $a < x$, $a \in \{0\} \cup E$, and $(a, x) \cap E = \varnothing$. Thus if we let s be the starting date $\nu(P, a + 1)$, then since D prices P on x, $s \le x$ by Proposition 7.5. Since $a < s$ it follows that $[s, x) \cap E = \varnothing$.

Therefore, since s is a starting date for P, $e \in E$, and $[s, e) = [s, x) \cup [x, e)$ contains no ending dates for P, $[s, e)$ is a tracking period for P containing x.

Miscellaneous Functions

Define the function **prc**: $\mathcal{P} \times \mathbb{M} \mapsto \mathbb{R}$ as follows: Fix any $(P, d) \in \mathcal{P} \times \mathbb{M}$.

If d does not belong to a tracking period for P, then set $\text{prc}(P, d) = 0$.

If d belongs to a tracking period $[s, e)$ for P, then by Proposition 2.5, $\pi(P, d) \ge s$, and thus by Proposition 2.1, D prices P on $\pi(P, d)$. Set $\text{prc}(P, d)$ equal to D’s price for P on $\pi(P, d)$.

Note that $\text{prc}(P, d) > 0$ if and only if d belongs to a tracking period for P. The condition does *not* imply that D prices P on d. Thus, contrary to what its name might suggest, the “pricing function” can be used to test whether d belongs to a tracking period for P, but not to test whether D prices P on d.

Before defining the next function, we need one more axiom:

Axiom 6: The pair of fields (permno, shrsdt) forms a key for D’s Shares table.

With this axiom, we may define the function **shrout**: $\mathcal{P} \times \mathbb{M} \mapsto \mathbb{Z}$, **D’s shares outstanding function**, as follows: Fix any $(P, d) \in \mathcal{P} \times \mathbb{M}$.

If d does not belong to a tracking period for P then set $\text{shrout}(P, d) = 0$.

If d belongs to a tracking period $[s, e)$ for P, then let $S = \{x \in [s, e) \mid$ there exists a record R in D’s Shares table with R.permno = P and R.shrsdt = x$\}$.

If $[s, d] \cap S \ne \varnothing$ then set shrout(P, d) = R.shrout where R is the record in D’s Shares table with R.permno = P and R.shrsdt = $\max([s, d] \cap S)$.

If $[s, d] \cap S = \varnothing$ but $(d, e) \cap S \neq \varnothing$ then set shrout(P, d) = R.shrout, where R is the record in D's Shares table with R.permno = P and R.shrsdt = $\min((d, e) \cap S)$.

Otherwise, set shrout(P, d) = 0.

Next, define the function **mval**: $\mathcal{P} \times \mathbb{M} \mapsto \mathbb{R}$, **D's market value function**, by

$$\text{mval}(P, d) = \lfloor \text{shrout}(P, d) \cdot |\text{prc}(P, d)| \rfloor + \frac{P}{100{,}000}$$

for all $(P, d) \in \mathcal{P} \times \mathbb{M}$, where the mapping $\lfloor\ \rfloor: \mathbb{R} \mapsto \mathbb{Z}$ is the so-called "greatest integer function" defined by $\lfloor x \rfloor = \max(\{n \in \mathbb{Z} \mid n \leq x\})$ for all $x \in \mathbb{R}$. Note that $x - \lfloor x \rfloor$ is x's "remainder" upon division by 1, which we use in the proof of the following:

Proposition 12: If (P_1, d_1) and (P_2, d_2) are elements of $\mathcal{P} \times \mathbb{M}$ where $\text{mval}(P_1, d_1) = \text{mval}(P_2, d_2)$, then $P_1 = P_2$.

> *Proof*: Observe that $0 \leq P / 100{,}000 < 1$ for all $P \in \mathcal{P}$. Thus $P_1 / 100{,}000 = \text{mval}(P_1, d_1) - \lfloor \text{mval}(P_1, d_1) \rfloor = \text{mval}(P_2, d_2) - \lfloor \text{mval}(P_2, d_2) \rfloor = P_2 / 100{,}000$, which implies $P_1 = P_2$.

This proposition implies that any collection of permno's in D has a unique member with the largest market value on a given date, if only because ties in true market value are broken by sorting on permno.

Define the function **mcap**: $\mathcal{C} \times \mathbb{M} \mapsto \mathbb{R}$, **D's market capitalization function**, by

$$\text{mcap}(C, d) = \sum_{\substack{Q \in \mathcal{P}, \\ \text{permco}(Q) = C}} \text{mval}(Q, d)$$

for all $(C, d) \in \mathcal{C} \times \mathbb{M}$.

In other words, mcap(C, d) is the sum of the market values on d over all active permno's in D whose permco is equal to C (the contribution to the sum from permno's that are not trading will be insignificant). The summation is required to account for all of C's various share classes that might be trading on d in order to get a get a figure that comes closest to C's full market capitalization in the usual sense. Note that if C is a foreign company whose only security traded in the US is an ADR, mcap will greatly understate the full market capitalization of C unless the CRSP data is supplemented appropriately.

Define the mapping **primno**: $\mathcal{C} \times \mathbb{M} \mapsto \mathcal{P}$, **D's primary permno function**, as follows: Fix any $(C, d) \in \mathcal{C} \times \mathbb{M}$. Let $A = \{P \in \mathcal{P} \mid \text{permco}(P) = C \text{ and D prices P on } \nu(P, d)\}$.

> If $A = \varnothing$ then define primno(C, d) = 0.
>
> Otherwise let $m = \min(\{\nu(P, d) \mid P \in A\})$ and let $B = \{P \in A \mid \nu(P, d) = m\}$. Define primno(C, d) to be the element of B with the largest market value on m, which is unique by Proposition 12.

Note that by definition, D always prices primno(C, d) on $\nu(\text{primno}(C, d), d)$.

The Branches Table

In this section we will add a new table called Branches to D in five chronological steps.

Step 1: Create a table in D called Branches by selecting all records from D's Distributions table where (exdt – 1) belongs to a tracking period for permno, taking only the fields

permno	
distcd	(for "distribution code")
divamt	(for "dividend amount")
facpr	(for "factor to adjust price")
exdt	(for "ex-distribution date")
acperm	(for "acquiring permno")
accomp	(for "acquiring permco")

and adding the following six new fields,

basdt	(for "basis date")
rsmdt	(for "resumption date")
rsmval	(for "resumption value")
bpermno	(for "branch permno")
bpermco	(for "branch permco")
brsmdt	(for "branch resumption date")

Step 2: For each record R in Branches such that $6000 \le R.distcd < 7000$ and $-1 < R.facpr < 0$, make the following modifications to Branches:

1. Add exactly one new record, R', to Branches that is identical to R.
2. Reset R.distcd = R.distcd + 2000.
3. Reset R'.distcd = R'.distcd + 3000.
4. Reset R.divamt = (R.divamt)·(–R.facpr).
5. Reset R'.divamt = 0.
6. Reset R.facpr = 0.
7. Reset R'.acperm = R'.accomp = 0.

Note that after these modifications all records in Branches retain the property that (exdt – 1) belongs to a tracking period for permno.

Step 3: For each record R in Branches, set

$R.basdt = \pi(R.permno, R.exdt - 1)$
$R.rsmdt = \nu(R.permno, R.exdt)$

Note that both of these values depend only on R.permno and R.exdt.

Step 4: For each record R in Branches, set the values of R.bpermno and R.bpermco using the following procedure:

1. Let P = R.permno, let C = permco(P), let d = R.exdt, and let [s, e) be the tracking period for P containing d – 1.

2. If d = 99991231 then reset R.bpermno = R.bpermco = 0 and exit the procedure. Otherwise reset R.bpermno = R.acperm, reset R.bpermco = R.accomp, and proceed to Step 4.3.

3. If R.bpermno ≠ 0, R.bpermno ∈ $\mathcal{P}$, and D prices R.bpermno on ν(R.bpermno, d) then reset R.bpermco = permco(R.bpermno) /*in case it isn't already*/ and exit the procedure. Otherwise proceed to Step 4.4.

4. Let Q = primno(R.bpermco, d). If Q ≠ 0 and either R.bpermco ≠ C or ν(Q, d) < ν(P, d) then reset R.bpermno = Q and exit the procedure. Otherwise proceed to Step 4.5.

5. If D prices P on ν(P, d), then reset R.bpermno = P, reset R.bpermco = C, and exit the procedure. Otherwise proceed to Step 4.6.

6. Let S be the record in D's Delistings table with permno = P and enddt = e. If R.bpermno ≠ S.nwperm or R.bpermco ≠ S.nwcomp, then reset R.bpermno = S.nwperm, reset R.bpermco = S.nwcomp, and return to Step 4.3. Otherwise proceed to Step 4.7.

7. If d belongs to a tracking period for some P' where permco(P') = C and D prices P' on ν(P', d), then reset R.bpermno = primno(C, d) and R.bpermco = C. Otherwise reset R.bpermno = R.bpermco = 0.

Step 5: For each record R in D's Branches set R.brsmdt = ν(R.bpermno, R.exdt).

The values in the field rsmval will be set later.

Proposition 13: For each record R in D's Branches table:
1. R.exdt – 1 and R.rsmdt – 1 belong to a common tracking period for R.permno.
2. Either R.exdt = 99991231 or D prices R.bpermno on R.brsmdt.

Proof: Let R be any record in D's Branches table and let P = R.permno and d = R.exdt.

That R.exdt – 1 = d – 1 belongs to some tracking period [s, e) for P is guaranteed in Steps 1-2 in the construction of D's Branches table. Since R.rsmdt = ν(P, d) and d ≤ e, it follows from Propositions 7.2 and 7.5 that d ≤ R.rsmdt ≤ e, which implies R.rsmdt – 1 ∈ [s, e), thus proving (1).

We will prove (2) by examining each of the exit points for the procedure in Step 4, which are located in Steps 4.2, 4.3, 4.4, 4.5, and 4.7. For the one in 4.2, there is nothing to prove because R.exdt is assumed to be 99991231.

For the other exits, note that by Step 5, R.brsmdt = ν(R.bpermno, d). In order to exit at 4.3, it is explicitly required that D price R.bpermno on ν(R.bpermno, d). For exits at 4.4, (2) is implied by the comment immediately following the definition of the primno mapping. In order to exit at 4.5, D must price P on ν(P, d), but R.bpermno is reset to P, so (2) holds. In 4.7, there are two possible exits: for the first, (2) follows, once again, from the comment immediately following the definition of the primno mapping. In the second, R.bpermno is set to 0, which D prices on every market date, including R.brsmdt = $\nu(0, \text{R.exdt})$ = R.exdt < 99991231.

Accumulation Periods

Fix any P ∈ $\mathcal{P}$. Let L be the set of P's delisting dates, E the set of P's ending dates, and T the set of P's trailer periods. Let λ, ε, and τ be defined according to Proposition 5.

An extended market date b is defined to be a **branch date** for P if either

1. P = 0 and b = 99991231 or
2. there exists a record R in D's Branches table with R.permno = P, R.rsmdt = b, and R.bpermno ≠ P.

Note that since R.rsmdt = ν(R.permno, R.exdt) for any record R in D's Branches table, it follows from Proposition 7.11 that ν(P, b) = b for any branch date b for P.

Proposition 14: Every ending date for P is a branch date for P. //Previously Propositoin 15.

Proof: The proposition is trivial if P = 0 because 99991231 is its only ending date, so assume P ≠ 0 and let e be any ending date for P. By Axiom 5, we may let S be the unique record in D's Delistings table with S.permno = P and S.dlstdt = $\varepsilon^{-1}(e)$ ∈ L. Each case handled in the procedure setting the value of S.enddt ensures the existence of a record, R, in D's Distributions table with R.permno = P and R.exdt = S.enddt = e. Since e − 1 belongs to a tracking period for P, Step 1 in the construction of D's Branches table ensures that a record R' with R'.permno = P and R'.exdt = e appears in D's Branches table. By Step 5 of the construction of D's Branches table, R'.rsmdt = ν(R'.permno, R'.exdt) = ν(P, e) = e. To complete the proof we must prove that R'.bpermno ≠ P.

If e = 99991231 then Step 4.2 of the construction of D's Branches table implies R'.bpermno = 0 ≠ P, so assume e < 99991231. Since e is an ending date for P, D does not price P on ν(P, e) = e. But by Proposition 13.2, D prices R'.bpermno on R'.brsmdt = ν(R'.bpermno, e). It follows that R'.bpermno ≠ P.

Define the market interval [a, b) to be an **accumulation period** for P if

1. a is a branch date or starting date for P,
2. b is a branch date for P, and
3. (a, b) contains no branch dates or starting dates for P.

Note that since min($\mathbb{M}$) is 0's only starting date and 99991231 is 0's only branch date, [min($\mathbb{M}$), 99991231) = $\mathbb{M}$ is 0's only accumulation period.

Proposition 15: Distinct accumulation periods for P are disjoint. //Previously Propositoin 16

Suppose [a, b) and [a', b') are accumulation periods for P with nonempty intersection. Then by Lemma 1, g = max(a, a') is a branch date or starting date for P contained in both [a, b) and [a', b'). Since (a, b) contains no branch dates or starting dates for P, we must have g = a. But (a', b') contains no branch dates or starting dates for P, so g = a'. Hence a = a'.

Similarly, $h = \min(b, b')$ is a branch date for P contained in $(a, b] \cap (a', b']$. Since (a, b) contains no branch dates for P, $h = b$, and since (a', b') contains no branch dates for P, $h = b'$. Thus $b = b'$.

Proposition 16: If d belongs to some tracking period [s, e) for P (which is always the case if D prices P on d), then there exists an accumulation period [a, b) for P (unique by Proposition 15) such that $d \in [a, b) \subset [s, e)$. //Previously Propositoin 17

Proof: Let $F = \{d \in \mathbb{M} \mid d$ is a branch date or a starting date for P$\}$, $A = \{x \in F \mid x \leq d\}$ and $B = \{x \in F \mid x > d\}$. Since $s \in A$ and, by Proposition 14, $e \in B$, neither is empty, so let $a = \max(A)$ and $b = \min(B)$. Then $a, b \in F$ and $s \leq a \leq d < b \leq e$, implying $d \in [a, b) \subset [s, e)$. If $b = e$ then b is a branch date for P. If $b < e$ then b belongs to [s, e), and thus is not a starting date for P by Corollary 9. Thus since $b \in F$, b must be a branch date for P. Since $a \in F$, a is either a starting date or a branch date for P. By construction, (a, b) contains no elements of F, and thus contains no branch dates or starting dates for P. Thus [a, b) is an accumulation period for P.

The last two propositions imply that each tracking period for P is partitioned by accumulation periods for P. The next proposition implies that every accumulation period for P is a member of such a partition:

Proposition 17: Every accumulation period for P is contained within a tracking period for P.
//Previously Propositoin 18

Proof: Let [a, b) be an accumulation period for P. Since b is a branch date for P, there exists a record R in D's Branches table with R.permno = P, R.rsmdt = b, and R.bpermno ≠ P. By Proposition 13.1, $b - 1$ belongs to a tracking period [s, e) for P. Thus by the last proposition, there exists an accumulation period [a', b') for P such that $b - 1 \in [a', b') \subset [s, e)$.

Since $b - 1 \in [a, b) \cap [a', b')$, Proposition 15 implies that $[a, b) = [a', b')$. Thus $[a, b) \subset [s, e)$.

Accumulators

An ordered pair $\alpha = (P, [a, b))$ is an **accumulator** in D if $P \in \mathcal{P}$ and [a, b) is an accumulation period for P. Let $\mathcal{A}$ be the set of all accumulators in D. Note that $(0, \mathbb{M}) \in \mathcal{A}$.

For each $\alpha = (P, [a, b)) \in \mathcal{A}$, define **α's records in Branches,** abbreviated $\mathbf{B_\alpha}$, to be the set of all records in D's Branches table where permno = P and rsmdt $\in (a, b]$. For any $d \in \mathbb{M}^*$, define $\mathbf{B_\alpha(d)} = \{R \in B_\alpha \mid$ R.rsmdt = d$\}$.

Proposition 18: Suppose $\alpha = (P, [a, b)) \in \mathcal{A}$. If $R \in B_\alpha$ and R.bpermno ≠ P then $R \in B_\alpha(b)$.
//Previously Propositoin 19

Proof: By the definition of B_α, R.rsmdt $\in (a, b]$. Since R.rsmdt is a branch date for P and (a, b) contains no branch dates for P, it follows that R.rsmdt = b and thus that $R \in B_\alpha(b)$.

Proposition 19: The collection $\{B_\alpha\}_{\alpha \in \mathcal{A}}$ partitions the set of records in D's Branches table. //Previously Propositoin 20

Proof: Given any record R in D's Branches table, R.rsmdt – 1 belongs to a tracking period for R.permno by Proposition 13.1, and thus to some accumulation period [a, b) for R.permno by Proposition 16. Thus R.rsmdt $\in$ (a, b] and α = (R.permno, [a, b)) $\in \mathcal{A}$, so R $\in B_\alpha$. If R also belongs to $B_{\alpha'}$ where α' = (P', [a', b')), then P' = R.permno = P and R.rsmdt $\in$ [a, b) $\cap$ [a', b'), which by Proposition 15 implies that [a, b) = [a', b'). Thus $\alpha = \alpha'$, which completes the proof.

We are now finally ready to define the concept from which D's Branches table takes its name: If α = (P, [a, b)) and α' = (P', [a', b')) are accumulators in D, we will say that **α branches to α'**, abbreviated by **$\alpha \to \alpha'$**, if either

1. P = P' and b = a' or
2. P $\neq$ P' and there exists a record R $\in B_\alpha$(b) with R.bpermno = P' and R.brsmdt $\in$ [a', b').

Proposition 20: Suppose α = (P, [a, b)) $\in \mathcal{A}$. If P $\neq$ 0 then $B_\alpha(b) \neq \varnothing$, and if b < 99991231 then for each R $\in B_\alpha$(b) there exists a unique α' = (P', [a', b')) $\in \mathcal{A}$ such that P' = R.bpermno and $\alpha \to \alpha'$. //Previously Propositoin 21.

Proof: If P $\neq$ 0 then since b is a branch date for P, by definition there exists at least one record, R, in D's Branches table with R.permno = P and R.rsmdt = b, meaning R $\in B_\alpha(b) \neq \varnothing$.

Assume b < 99991231 and let R be any element of B_α(b). Let P' = R.bpermno. Since R.exdt $\leq$ ν(P, R.exdt) = R.rsmdt = b < 99991231, Proposition 13.2 implies that D prices P' on R.brsmdt.

Assume P = P'. Then D prices P on R.brsmdt = ν(P', R.exdt) = ν(P, R.exdt) = R.rsmdt = b, meaning b is not an ending date for P. Therefore since b – 1 belongs to a tracking period for P by Proposition 13.1, so does b. Proposition 16 thus implies that b belongs to an accumulation period [a', b') for P. Since b is a branch date for P and (a', b') contains no branch dates for P, we must conclude that b = a'. If we let α' = (P, [b, b')) then it follows that $\alpha \to \alpha'$.

Assume P' $\neq$ P. Since D prices P' on R.brsmdt, R.brsmdt belongs to an accumulation period [a', b') for P' by Proposition 16. If we let α' = (P', [a', b')) then it follows that $\alpha \to \alpha'$.

As for the uniqueness of α', note that if $\alpha \to \alpha''$ = (P'', [a'', b'') where P'' = R.bpermno, then P'' = P'. Since [a', b') and [a'', b'') either both contain b (if P = P') or both contain R.brsmdt (if P' $\neq$ P), Proposition 15 implies [a', b') = [a'', b''), and thus $\alpha' = \alpha''$.

If S $\subset \mathcal{A}$, then we will say that **α branches to S**, abbreviated **$\alpha \to$ S**, if $\{\alpha' \in \mathcal{A} \mid \alpha \to \alpha'\} \subset$ S.

That is, α branches to S if all accumulators to which α branches are contained in S. Note that the set $\{\alpha' \in \mathcal{A} \mid \alpha \to \alpha'\}$ could very well be empty (namely, whenever, and only when, b = 99991231), in which case α trivially branches to *any* subset of $\mathcal{A}$, even though α branches to no particular accumulator.

Define $\mathcal{A}_0 = \{(0, \mathbb{M})\}$. For each n > 0, define $\mathcal{A}_n$ inductively by $\mathcal{A}_n = \{\alpha \in \mathcal{A} \mid \alpha \to \mathcal{A}_{n-1}\}$. Observe that $\mathcal{A}_1$ includes all accumulators that branch to no accumulators in D, in addition to all those that branch to $(0, \mathbb{M})$. That is, if $\{\alpha \in \mathcal{A} \mid \alpha_1 \to \alpha\} = \varnothing$ then $\alpha_1 \in \mathcal{A}_1$.

Proposition 21: $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ for all $n \geq 0$. //Previously Propositoin 22.

Proof: First observe that $\{\alpha \in \mathcal{A} \mid (0, \mathbb{M}) \to \alpha\} = \{(0, \mathbb{M})\} \subset \{(0, \mathbb{M})\} = \mathcal{A}_0$, so $(0, \mathbb{M}) \in \mathcal{A}_1$. Thus $\mathcal{A}_0 \subset \mathcal{A}_1$.

Assume that $\mathcal{A}_k \subset \mathcal{A}_{k+1}$ for some $k \geq 0$ and fix any $\alpha' \in \mathcal{A}_{k+1}$. Then by definition, $\alpha' \to \mathcal{A}_k$, meaning $\{\alpha \in \mathcal{A} \mid \alpha' \to \alpha\} \subset \mathcal{A}_k$. But we are assuming $\mathcal{A}_k \subset \mathcal{A}_{k+1}$, so $\{\alpha \in \mathcal{A} \mid \alpha' \to \alpha\} \subset \mathcal{A}_{k+1}$. Thus $\alpha' \to \mathcal{A}_{k+1}$, and hence $\alpha' \in \mathcal{A}_{k+2}$. Since $\alpha' \in \mathcal{A}_{k+1}$ was arbitrary, we have proved that $\mathcal{A}_{k+1} \subset \mathcal{A}_{k+2}$. The proposition now follows by induction.

This proposition implies that $\{\mathcal{A}_n\}_{n=1}^{\infty}$ is a monotonically increasing sequence of subsets of $\mathcal{A}$. Since $\mathcal{A}$ is finite, the sequence cannot be strictly increasing. Thus there must exist some (infinitely many, in fact) $n \in \mathbb{Z}^+$ such that $\mathcal{A}_n = \mathcal{A}_{n+1}$. Let $N = \min(\{n \in \mathbb{Z}^+ \mid \mathcal{A}_n = \mathcal{A}_{n+1}\})$.

The obvious question is, does $\mathcal{A} = \mathcal{A}_N$? As it turns out, the answer is “yes” (Proposition 23). In order to prove this important claim critical to defining total returns for accumulators, we need to take a close look at the branching relationship:

Define the mapping $\psi: \mathcal{A} \times \mathcal{A} \mapsto \mathbb{M}^*$ as follows:

If $\alpha = (P, [a, b)) \not\to \alpha' = (P', [a', b'))$ then define $\psi(\alpha, \alpha') = 0$.

If $\alpha = (P, [a, b)) \to \alpha' = (P', [a', b'))$ and $P = P'$ then define $\psi(\alpha, \alpha') = b$.

If $\alpha = (P, [a, b)) \to \alpha' = (P', [a', b'))$ and $P \neq P'$ then define $\psi(\alpha, \alpha') = \max(\{R.\text{exdt} \mid R \in B_\alpha, R.\text{bpermno} = P' \text{ and } R.\text{brsmdt} \in [a', b')\})$.

Note that the set to which the max function is applied in the case where $P \neq P'$ is nonempty according to the definition of $\alpha \to \alpha'$. Also note that the set might contain more than one record (such as when an acquiring company distributes liquidation payments in the form of a series of spin-off’s during the acquired stock’s trailer period)—since we must define $\psi(\alpha, \alpha')$ to be something, we have arbitrarily set it equal to the maximum value of exdt over such records.

Lemma 22: If $\alpha \to \alpha' \to \alpha''$ then $\psi(\alpha, \alpha') < \psi(\alpha', \alpha'')$. //Previously Lemma 23.

Proof: Suppose $\alpha = (P, [a, b))$, $\alpha' = (P', [a', b'))$, $\alpha'' = (P'', [a'', b''))$, and $\alpha \to \alpha' \to \alpha''$. We will prove the lemma case by case:

First assume $P = P'$. Then $\psi(\alpha, \alpha') = b = a'$.

If $P' = P''$ then $a' < b' = \psi(\alpha', \alpha'')$, and thus the conclusion follows.

If $P' \neq P''$ then fix any $R' \in B_{\alpha'}(b')$ such that R'.bpermno = P'', R'.brsmdt $\in$ [a'', b'') and R'.exdt = $\psi(\alpha', \alpha'')$. Since a' is a branch date for P' we have $a' = \nu(P', a')$. For any $x \leq a'$, Proposition 7.5 implies $\nu(P', x) \leq a'$. But since $\nu(P', R'.exdt) = R'.rsmdt = b' > a'$, it follows that R'.exdt > a'. Thus the conclusion follows.

Next assume $P \neq P'$. Fix any $R \in B_\alpha(b)$ such that R.bpermno = P', R.brsmdt $\in$ [a', b') and R.exdt = $\psi(\alpha, \alpha')$. Then $\psi(\alpha, \alpha') = R.exdt \leq \nu(P', R.exdt) = R.brsmdt < b'$.

If $P' = P''$ then $\psi(\alpha', \alpha'') = b'$, and thus the conclusion follows.

If $P' \neq P''$ then fix any $R' \in B_{\alpha'}(b')$ such that R'.bpermno = P'', R'.brsmdt $\in$ [a'', b'') and R'.exdt = $\psi(\alpha', \alpha'')$. For any $x \geq R'.exdt$, Proposition 7.5 implies $\nu(P', x) \geq \nu(P', R'.exdt) = b'$. But since $\nu(P', R.exdt) = R.brsmdt < b'$, it follows that R.exdt < R'.exdt, from which the conclusion follows.

Proposition 23: $\mathcal{A} = \mathcal{A}_N$. //Previously Propositoin 24.

Proof: Assume, to obtain a contradiction, that there exists some $\alpha_0 \in \mathcal{A}\backslash\mathcal{A}_N$. Furthermore, suppose we have found a finite sequence $\{\alpha_i\} \subset \mathcal{A}$ (i = 0, 1, 2, ... , k) such that $\alpha_i \notin \mathcal{A}_N$ for all $0 \leq i \leq k$ and $\alpha_0 \rightarrow \alpha_1 \rightarrow \cdots \rightarrow \alpha_k$. This can certainly be done for k = 1, for otherwise $\alpha_0 \in \mathcal{A}_1 \subset \mathcal{A}_N$, contradicting our assumption.

Since $\mathcal{A}_N = \mathcal{A}_{N+1}$, $\alpha_k \notin \mathcal{A}_{N+1}$. The latter non-containment implies that there exists some $\alpha_{k+1} \notin \mathcal{A}_N$ such that $\alpha_k \rightarrow \alpha_{k+1}$. Thus by induction there exists an infinite sequence $\{\alpha_i\} \subset \mathcal{A}$ such that $\alpha_i \notin \mathcal{A}_N$ for all i and $\alpha_0 \rightarrow \alpha_1 \rightarrow \alpha_2 \rightarrow \cdots$. But by Lemma 22 it follows that $\psi(\alpha_0, \alpha_1) < \psi(\alpha_1, \alpha_2) < \psi(\alpha_2, \alpha_3) < \cdots$, which is impossible, because $\mathbb{M}^*$ is finite. This contradiction implies that $\mathcal{A}\backslash\mathcal{A}_N = \varnothing$, that is, that $\mathcal{A} = \mathcal{A}_N$.

Propositions 22 and 24 imply that $\{\mathcal{A}_0, \mathcal{A}_1\backslash\mathcal{A}_0, \mathcal{A}_2\backslash\mathcal{A}_1, \ldots, \mathcal{A}_N\backslash\mathcal{A}_{N-1}\}$ forms a partition of $\mathcal{A}$, a fact that will be crucial to our definition of total returns on accumulators. But first we need a couple more functions:

Split-Adjustment Factors

Fix any $P \in \mathcal{P}$. Define the mapping $\mathbf{s_P}$**:** $\mathbb{M} \rightarrow \mathbb{R}$ as follows: Fix any $d \in \mathbb{M}$. Let U be the set of all records in D's Branches table with permno = P, exdt = d, divamt = 0, facpr ≠ 0, facpr ≠ –1, and distcd ≥ 5000. If $U = \varnothing$ then set $s_P(d) = 0$. Otherwise, set

$$s_P(d) = \sum_{R \in U} R.\text{facpr}\,.$$

Next define the mapping $\mathbf{S_P}$**:** $\mathbb{M} \rightarrow \mathbb{R}$ by

$$S_P(d) = \prod_{x \in \mathbb{M},\, x \leq d} (1 + s_P(x))\,.$$

for all $d \in \mathbb{M}$.

Total Return Calculations: Accumulators

For each $\alpha \in \mathcal{A}$ define the functions $\mathbf{Ret_{\alpha}}$**:** $\mathbb{M} \to \mathbb{R}$ and $\mathbf{Liq_{\alpha}}$**:** $\mathbb{M} \to$ **{0, 1}** and set the values in rsmval for all records in B_{α} as follows:

Let N be the constant referred to in Proposition 23.

For each $d \in \mathbb{M}$ with $d > \min(\mathbb{M})$, set $Ret_{(0, M)}(d) = prc(0, d)/prc(0, d - 1)$ and set $Liq_{(0, M)}(d) = 1$. Set $Ret_{(0, M)}(\min(\mathbb{M})) = Liq_{(0, M)}(\min(\mathbb{M})) = 1$. Thus $Ret_{\alpha}: \mathbb{M} \to \mathbb{R}$ and $Liq_{\alpha}: \mathbb{M} \to \{0, 1\}$ are defined for all $\alpha \in \mathcal{A}_0$. The only record in Branches with permno = 0 has rsmdt = 99991231, for which we set rsmval = prc(0, max($\mathbb{M}$)). Thus the values in rsmval are set for all records in B_{α} for all $\alpha \in \mathcal{A}_{0.}$

Assume, for some $0 < k \leq N$ and all $\alpha \in \mathcal{A}_{k-1}$, that $Ret_{\alpha}: \mathbb{M} \to \mathbb{R}$ and $Liq_{\alpha}: \mathbb{M} \to \{0, 1\}$ are defined and that rsmval has been set for all records in B_{α}. Since $N = \min(\{i \geq 0 \mid \mathcal{A}_i = \mathcal{A}_{i+1}\})$, we have $\mathcal{A}_{k-1} \neq \mathcal{A}_k$, which by Proposition 21 implies $\mathcal{A}_k \backslash \mathcal{A}_{k-1} \neq \emptyset$.

Fix any $\alpha = (P, [a, b)) \in \mathcal{A}_k \backslash \mathcal{A}_{k-1}$. Since $\alpha \notin \mathcal{A}_{k-1}$ we are free to define $Ret_{\alpha}: \mathbb{M} \to \mathbb{R}$ and $Liq_{\alpha}: \mathbb{M} \to \{0, 1\}$ without risk of contradiction. And by Proposition 19, we are free to set the values in rsmval for all records in B_{α} without risk of contradiction.

We will first set the values in rsmval for all records in B_{α}. To do so, let R be any record in D's Branches table with R.permno = P and R.rsmdt $\in$ (a, b]. Let P' = R.bpermno.

We first calculate the total return on P', represented by $R_{P'}$, from the time where the distribution's value, R.divamt, is defined (usually the close of R.exdt) to the close of b:

If b = 99991231 or P = P', then set $R_{P'} = 1$.

Otherwise, $P \neq P'$ implies $R \in B_{\alpha}(b)$ by Proposition 18, and thus since b < 99991231, we may let $\alpha' = (P', [a', b'))$ be the unique accumulator corresponding to R by Proposition 20. Since $\alpha \in \mathcal{A}_k$ and $\alpha \to \alpha'$, it follows by definition that $\alpha' \in \mathcal{A}_{k-1}$. Therefore by our induction hypothesis, $Ret_{\alpha'}: \mathbb{M} \to \mathbb{R}$ is defined. We may therefore set

$$R_{P'} = \begin{cases} Ret_{\alpha'}(b) & \text{if R.distcd} < 0, \\ \displaystyle\prod_{x = \text{R.brsmdt} + 1}^{b} Ret_{\alpha'}(x) & \text{if R.distcd} > 0 \text{ and R.brsmdt} < b\text{, and} \\ 1 & \text{otherwise.} \end{cases}$$

With $R_{P'}$ defined, set the value of R.rsmval equal to

$$\frac{S_P(R.exdt-1)}{S_P(R.basdt)} \cdot R_{P'} \cdot \begin{cases} R.facpr \cdot |prc(P, R.rsmdt)| & \text{if } R.distcd \geq 3000,\ R.facpr \neq -1,\ R.facpr \neq 0, \\ & \text{and D prices P on } R.rsmdt, \text{ and} \\ R.divamt & \text{otherwise.} \end{cases}$$

With the values in rsmval set for all records in B_α, define $Ret_\alpha: \mathbb{M} \to \mathbb{R}$ and $Liq_\alpha: \mathbb{M} \to \{0, 1\}$ as follows:

By Proposition 17, there exists a tracking period [s, e) for P containing [a, b), meaning $b \in (s, e]$.

If D prices P on b, then $b \neq e$ and thus $b \in (s, e)$. By Proposition 16, there exists an accumulation period [d, c) for P such that $b \in [d, c)$. Since b is a branch date for P and (d, c) contains no branch dates for P, it follows that d = b. Thus if we let $\alpha_0 = (P, [b, c))$ then $\alpha \to \alpha_0$, which implies $\alpha_0 \in \mathcal{A}_{k-1}$. Let $v_0 = prc(P, b)$.

Otherwise, if D does not price P on b, then arbitrarily set $v_0 = 0$ and set $\alpha_0 = (0, \mathbb{M})$. In this case $\alpha_0 \in \mathcal{A}_0 \subset \mathcal{A}_{k-1}$.

In either case, $\alpha_0 \in \mathcal{A}_{k-1}$, which by our induction hypothesis implies that $Ret_{\alpha 0}: \mathbb{M} \to \mathbb{R}$ and $Liq_{\alpha 0}: \mathbb{M} \to \{0, 1\}$ are defined.

Let x be any market date in $\mathbb{M}$.

If x < a then set $Ret_\alpha(x) = 1$ and $Liq_\alpha(x) = 0$.

If x = a then set $Ret_\alpha(x) = 1$ and $Liq_\alpha(x) = 1$.

If $x \in (a, b)$ then $x - 1 \in [s, e)$, which implies $prc(P, x - 1) > 0$.

If D prices P on x, then set $Liq_\alpha(x) = 1$ and

$$Ret_\alpha(x) = \frac{prc(P, x) + \sum_{R \in B_\alpha(x)} R.rsmval}{prc(P, x-1)},$$

where the summation is zero if $B_\alpha(x)$ is empty.

Otherwise, if D does not price P on x then set $Ret_\alpha(x) = 1$ and $Liq_\alpha(x) = 0$.

If $x \geq b$, then $b \leq x < 99991231$. Since $\alpha \notin \mathcal{A}_0$, we must have $P \neq 0$. Thus by Proposition 20, we may let $R_1, R_2, \ldots, R_n$ be the distinct elements of $B_\alpha(b)$ and let α_i be the unique accumulator corresponding to R_i to which α branches for $1 \leq i \leq n$. Let let $v_i = R_i.rsmval$ for $1 \leq i \leq n$ and let $V = \Sigma_{0 \leq i \leq n} v_i$. (Recall that α_0 and v_0 were defined above.) By our induction hypothesis, $Ret_{\alpha i}: \mathbb{M} \to \mathbb{R}$ and $Liq_{\alpha i}: \mathbb{M} \to \{0, 1\}$ are defined for all $0 \leq i \leq n$. Set $Liq_\alpha(x) = \Pi_{0 \leq i \leq n} Liq_{\alpha i}(x)$.

If x = b then $b - 1 \in [s, e)$, so we may set $Ret_\alpha(x) = V/prc(P, b - 1)$.

If $x > b$ then set

$$\mathrm{Ret}_{\alpha}(x) = \frac{\sum_{i=0}^{n}\left[v_i \cdot \left(\prod_{d=b+1}^{x} \mathrm{Ret}_{\alpha i}(d) \right)\right]}{\sum_{i=0}^{n}\left[v_i \cdot \left(\prod_{d=b+1}^{x-1} \mathrm{Ret}_{\alpha i}(d) \right)\right]},$$

where

1. any products in the numerator with $x < b + 1$ are defined to be 1,
2. any products in the denominator with $x - 1 < b + 1$ are defined to be 1, and
3. if the denominator is zero, $\mathrm{Ret}_{\alpha}(x)$ is defined to be 1.

This defines $\mathrm{Ret}_{\alpha}(x)$ and $\mathrm{Liq}_{\alpha}(x)$ for all $x \in \mathbb{M}$.

Since $\alpha \in \mathcal{A}_k$ was arbitrary, $\mathrm{Ret}_{\alpha}: \mathbb{M} \to \mathbb{R}$ and $\mathrm{Liq}_{\alpha}: \mathbb{M} \to \{0, 1\}$ are defined for all $\alpha \in \mathcal{A}_k$. It follows by induction, then, that $\mathrm{Ret}_{\alpha}: \mathbb{M} \to \mathbb{R}$ and $\mathrm{Liq}_{\alpha}: \mathbb{M} \to \{0, 1\}$ are defined for all $\alpha \in A_k$ and all $0 \le k \le N$. By Proposition 23, we have thus defined $\mathrm{Ret}_{\alpha}: \mathbb{M} \to \mathbb{R}$ and $\mathrm{Liq}_{\alpha}: \mathbb{M} \to \{0, 1\}$ for all $\alpha \in \mathcal{A}$, as we intended.

The Roots Table

A permno $P \in \mathcal{P}$ is defined to be **active on d** if d belongs to a tracking period for P and D prices P on $\nu(P, d)$. Equivalently, P is active on d if and only if d belongs to a tracking period for P but does not belong to a trailer period for P.

Create a new table in D called Roots with the following nine fields:

permno	
permco	
rttype	(for "root type")
rtcd	(for "root code")
rtexdt	(for "root ex-date")
rtdt	(for "root date")
rtperm	(for "root permno")
rtcomp	(for "root permco")
rtsize	(for "root size")

Add records to D's Roots table according to the procedure given in the following six steps:

Step 1: For each record, S, in D's Branches table where

S.permno > 0,
S.divamt > 0,
S.exdt < 99991231,
S.bpermco > 0, and
S.bpermco ≠ permco(S.permno),

add records to D's Roots table as follows: Let P = S.bpermno and d = S.exdt. Let $U = \{P_1, P_2, \ldots P_n\}$ be the set of all permno's in $\mathcal{P}$ such that $\mathrm{permco}(P_i)$ = S.bpermco and either

1. P_i is active on d or
2. D prices P_i in (d, S.brsmdt]

for each i ∈ {1, 2,…n}. (Note that by Proposition 13.2, P itself qualifies as a member of U by (2).) For each i ∈ {1, 2,…n}, add exactly one new record, R_i, to D’s Roots table, setting

R_i.permno = P_i
R_i.permco = S.bpermco

$$R_i.\text{rttype} = \begin{cases} 1 & \text{if } P_i = P, \\ 2 & \text{otherwise.} \end{cases}$$

R_i.rtcd = S.distcd
R_i.rtexdt = d
R_i.rtdt = min({m ∈ $\mathbb{M}$ | d ≤ m and D prices P_i on m})
R_i.rtperm = S.permno
R_i.rtcomp = permco(S.permno)
R_i.rtsize = max({shrout(S.permno, d – 1) · S.divamt, 1})

Note that the set defining R_i.rtdt is nonempty since either P_i is active on d by (1), meaning D prices P_i on ν(P, d) ≥ d, or D prices P_i on some m ∈ (d, S.brsmdt] by (2).

Step 2: For each (P, d) ∈ $\mathcal{P} \times \mathbb{M}$ such that P is active on d – 1 and there eixists at least one record, S, in D’s Roots table where

S.permno = P,
S.rtdt = d, and
S.rttype ∈ {1, 2},

add exactly one new record, R, to D's Roots table, setting

R.permno = P
R.permco = permco(P)
R.rttype = 3
R.rtcd = 0
R.rtexdt = d
R.rtdt = d
R.rtperm = P
R.rtcomp = permco(P)
R.rtsize = mcap(permco(P), d – 1)

Step 3: For each record, S, in D’s Branches table where

permco(S.permno) = S.bpermco,
S.divamt > 0, and
S.brsmdt is a starting date for S.bpermno,

add records to D’s Roots table as follows: Let P = S.bpermno and d = S.exdt. Let U = {P_1, P_2, …P_n} be the set of all elements of $\mathcal{P}$ such that

1. $\text{permco}(P_i)$ = S.bpermco and
2. P_i has a starting date in [d, S.brsmdt]

for each $i \in \{1, 2, \ldots n\}$. (Note that U includes P itself.) For each $i \in \{1, 2, \ldots n\}$, add exactly one new record, R_i, to D's Roots table, setting

R_i.permno = P_i
R_i.permco = S.bpermco
$$R_i.\text{rttype} = \begin{cases} 4 & \text{if } P_i = P, \\ 5 & \text{otherwise.} \end{cases}$$
R_i.rtcd = S.distcd
R_i.rtexdt = d
R_i.rtdt = $\min(\{m \in \mathbb{M} \mid d \leq m \text{ and } m \text{ is a starting date for } P_i\})$
R_i.rtperm = S.permno
R_i.rtcomp = S.bpermco
R_i.rtsize = $\max(\{\text{shrout}(\text{S.permno}, d - 1) \cdot \text{S.divamt}, 1\})$

Note that condition (2) above ensures that the set defining R_i.rtdt is nonempty.

Step 4: For each $(P, s) \in \mathcal{P} \times \mathbb{M}$ such that

1. s is a starting date for P,
2. $\{Q \in \mathcal{P} \mid \text{permco}(Q) = \text{permco}(P) \text{ and } Q \text{ is active on } s - 1\} \neq \emptyset$, and
3. there exists no record, S, in D's Roots table with S.permno = P, S.rtdt = s, and S.rttype $\in \{4, 5\}$,

add exactly one new record, R, to D's Roots table as follows: Let A be the set defined in the second condition above. Let Q be the element of A with the largest market value on $s - 1$. Set

R.permno = P
R.permco = permco(P)
R.rttype = 6
R.rtcd = 0
R.rtexdt = $\pi(Q, s - 1) + 1$
R.rtdt = s
R.rtperm = Q
R.rtcomp = permco(P)
R.rtsize = $\text{mcap}(\text{permco}(P), s - 1)$

Step 5: For each $(P, s) \in \mathcal{P} \times \mathbb{M}$ such that s is a starting date for P and there exists no record, S, in D's Roots table where

S.permno = P,
S.rtdt = s, and
S.rttype $\in \{1, 2, 3, 4, 5, 6\}$,

add exactly one new record, R, to D's Roots table, setting

R.permno = P
R.permco = permco(P)
R.rttype = 7
R.rtcd = 0
R.rtexdt = s
R.rtdt = s
R.rtperm = 0
R.rtcomp = 0
R.rtsize = 1

Step 6: For each $(P, s) \in \mathcal{P} \times \mathbb{M}$ such that s is a starting date for P, let $U = \{R_1, R_2, \dots R_n\}$ be the set of all records in D's Roots table where

1. R_i.permno = P,
2. R_i.rtdt = s, and
3. R_i.rttype $\in \{4, 5\}$

for each $i \in \{1, 2, \dots n\}$. If $U \neq \varnothing$ then let

$$c = \frac{\text{mcap}(\text{permco}(P), s-1)}{\sum_{i=1}^{n} R_i.\text{rtsize}}$$

and, for each $i \in \{1, 2, \dots n\}$, multiply the value in R_i.rtsize by c.

Root Dates

A market date $d \in \mathbb{M}$ is defined to be a **root date** for $P \in \mathcal{P}$ if there exists a record, R, in D's Roots table with R.permno = P, R.rtdt = d, and, if specified as a **root date of type n** for P, R.rttype = n. In order to assist in proving statements about root dates, we summarize some basic properties of the Roots table, most of which are plainly evident from the table's construction procedure itself, in the following lemma:

Lemma 24: Let R be any record in D's Roots table. //Previously Lemma 25.

1. R.permco = permco(R.permno) and R.rtcomp = permco(rtperm).
2. If R.rttype $\in \{1, 2, 4, 5\}$, then there exists a record, S, in D's Branches table where
   S.permno = R.rtperm,
   S.distcd = R.rtcd,
   S.exdt = R.rtexdt, and
   S.bpermco = R.permco.
3. If R.rttype $\in \{1, 2\}$ then
   $0 <$ R.permco $\neq$ R.rtcomp and
   R.rtdt = min($\{m \in \mathbb{M}$ | R.rtexdt $\leq$ m and D prices R.permno on m$\}$).
4. If R.rttype $\in \{4, 5\}$ then
   $0 <$ R.permco = R.rtcomp and
   R.rtdt = min($\{m \in \mathbb{M}$ | R.rtexdt $\leq$ m and m is a starting date for R.permno$\}$).
5. If R.rttype = 3 then

$0 <$ R.permno = R.rtperm,
R.rtexdt = R.rtdt,
R.permno is active on R.rtdt – 1, and
R.rtdt is a root date of type 1 or 2 for R.permno.

6. If R.rttype $\in$ {4, 5, 6, 7} then R.rtdt is a starting date for R.permno.
7. If R.rttype = 6 then
   $0 <$ R.permco = R.rtcomp,
   R.rtexdt = π(R.rtperm, R.rtdt – 1) + 1
   R.rtperm is active on R.rtdt – 1, and
   R.rtdt is not a root date of type 4 or 5 for R.permno.
8. If R.rttype = 7 then
   R.rtcomp = 0,
   R.rtexdt = R.rtdt, and
   R.rtdt is not a root date of type 1, 2, 3, 4, 5, or 6 for R.permno.
9. R.rtexdt $\le$ R.rtdt.

Proof: As noted above, most of these properties are obvious from the construction of D's Roots table, and only a few require any deduction:

In (4), $0 <$ R.permco is implied by the fact that the record S in Step 3 of the construction must satisfy the condition that S.brsmdt be a starting date for S.bpermno. For by Proposition 13.1, S.exdt – 1 belongs to a tracking period for S.permno and is thus a member of $\mathbb{M}$. Thus S.brsmdt = ν(S.bpermno, S.exdt) $\ge$ S.exdt > S.exdt – 1 $\ge$ min($\mathbb{M}$). Since min($\mathbb{M}$) is 0's only starting date and S.brsmdt, a starting date for S.bpermno, is greater than min($\mathbb{M}$), it follows that S.bpermno $\ne$ 0, and therefore that R.permco = S.bpermco $\ne$ 0.

In (5), $0 <$ R.permno follows from the fact that R.rtdt is a root date of type 1 or 2 for R.permno because the latter implies, as we just proved, that R.permno $\ne$ 0.

In (7), $0 <$ R.permco follows from (6), the fact that R.rtdt – 1 $\in$ $\mathbb{M}$ (implying R.rtdt > min($\mathbb{M}$)), and the fact that min($\mathbb{M}$) is 0's only starting date in D.

(9) follows from (3), (4), (7) and (8).

Proposition 25: //Previously Propositoin 26.

1. If d is a starting date for P then d is a root date for P.
2. If d is a root date for P then D prices P on d.
3. min($\mathbb{M}$) is 0's only root date.

Proof: (1) is guaranteed by Step 5 in the construction of D's Roots table. (2) follows from Lemma 24.3-6. (3) follows from Lemma 24.3-5 and 25.7.

Evaluation Periods

Define the market interval [a, b) to be an **evaluation period** for P if

1. a is a root date for P,
2. b is a root date or an ending date for P, and

3. (a, b) contains no root dates or ending dates for P.

Note that since min($\mathbb{M}$) is 0's only root date and 99991231 is 0's only ending date, $\mathbb{M}$ = [min($\mathbb{M}$), 99991231) is 0's only evaluation period.

Proposition 26: Distinct evaluation periods for P are disjoint. //Previously Propositoin 27.

Proof: See the proofs of Propositions 10 and 16, which are similar.

Proposition 27: If d belongs to a tracking period [s, e) for P (which is always the case if D prices P on d), then there exists a unique evaluation period [a, b) for P such that d ∈ [a, b) ⊂ [s, e). //Previously Propositoin 28.

Proof: Let F = {x ∈ $\mathbb{M}$ | x is a root date or an ending date for P}, A = {x ∈ F | x ≤ d} and B = {x ∈ F | x > d}. Clearly e ∈ B, and since starting dates for P are root dates for P, s ∈ A. Thus neither A nor B is empty, so let a = max(A) and b = min(B). Then a, b ∈ F and s ≤ a ≤ d < b ≤ e, implying d ∈ [a, b) ⊂ [s, e). By the definition of a tracking period, [s, e) contains no ending dates for P, and thus in particular, a is not an ending date for P. But a ∈ F, so a must be a root date for P. Since b ∈ F, b is either a root date or an ending date for P. Finally, by construction, (a, b) contains no elements of F, and thus no root dates or ending dates for P. Thus [a, b) is an evaluation period for P. The uniqueness of [a, b) follows from Proposition 26.

The last two propositions imply that each tracking period for P is partitioned by evaluation periods for P. The next proposition implies that every evaluation period for P is a member of such a partition:

Proposition 28: Every evaluation period for P is contained in a tracking period for P. //Previously Propositoin 29.

Proof: Let [a, b) be an evaluation period for P. By the definition of a root date for P, D prices P on a, and thus by Proposition 11, a belongs to a tracking period [s, e) for P. Thus by the last proposition, there exists an evaluation period [a', b') for P such that a ∈ [a', b') ⊂ [s, e). Since a ∈ [a, b) ∩ [a', b'), Proposition 24 implies that [a, b) = [a', b'). Thus [a, b) ⊂ [s, e).

Evaluators

An ordered pair ε = (P, [a, b)) is defined to be an **evaluator** in D if P ∈ $\mathcal{P}$ and [a, b) is an evaluation period for P. Let $\mathcal{E}$ be the set of all evaluators in D. Note that (0, $\mathbb{M}$) ∈ $\mathcal{E}$.

If ε = (P, [a, b)) and ε' = (P', [a', b')) are evaluators in D with a > min($\mathbb{M}$), we will say that **ε is rooted in ε'**, abbreviated by **ε' → ε**, if there exists a record, R, in D's Roots table where R.permno = P, R.rtperm = P', R.rtdt = a, and either

1. R.rttype ≠ 3 and R.rtexdt – 1 ∈ [a', b') or
2. R.rttype = 3 and b' = a.

Note that the abbreviation used here is identical to that used for the branching relationship. We will rely on context for distinguishing between branch and root relationships.

Proposition 29: Suppose $\varepsilon = (P, [a, b)) \in \mathcal{E}$ with $a > \min(\mathbb{M})$. For each record R in D's Roots table with R.permno = P and R.rtdt = a (of which there is at least one because a is a root date for P), there exists a unique evaluator $\varepsilon' = (P', [a', b'))$ such that R.rtperm = P' and $\varepsilon' \rightarrow \varepsilon$. //Previously Propositoin 30.

> *Proof*: Fix such a record, R, in D's Roots table and let P' = R.rtperm and d = R.rtexdt.
>
> If R.rttype $\in \{1, 2, 4, 5\}$ then by Lemma 24.2, Proposition 13.1, and Proposition 27, it follows that $d - 1$ belongs to a unique evaluation period [a', b') for P'. If we let $\varepsilon' = (P', [a', b'))$ then $\varepsilon' \rightarrow \varepsilon$, as intended.
>
> If R.rttype = 3 then by Lemma 24.5, P' = P and a = d. Also by Lemma 24.5, P is active on $a - 1$ and thus $a - 1$ belongs to a tracking period for P. Thus by Proposition 27, $a - 1$ belongs to an evaluation period [a', b') for P, which implies $a \in (a', b']$. Since (a', b') contains no root dates for P, we must conclude a = b'. If we let $\varepsilon' = (P, [a', b'))$ then $\varepsilon' \rightarrow \varepsilon$, as intended.
>
> If R.rttype = 6 then by Lemma 24.6, P' is active on $a - 1$ and $d = \pi(P', a - 1) + 1$. Thus $a - 1$ belongs to a tracking period [s, e) for P', and by Proposition 2.5, $s \leq \pi(P', a - 1) = d - 1$. By Proposition 2.2, $d - 1 = \pi(P', a - 1) \leq a - 1 < e$. Hence $d - 1 \in [s, e)$, so by Proposition 27, $d - 1$ belongs to an evaluation period [a', b') for P'. If we let $\varepsilon' = (P', [a', b'))$ then $\varepsilon' \rightarrow \varepsilon$, as intended.
>
> If R.rttype = 7, then Lemma 24.8 implies P' = 0 and d = a. Since $a > \min(\mathbb{M})$, $a - 1 \in \mathbb{M}$. If we let $\varepsilon' = (0, \mathbb{M})$ then $\varepsilon' \rightarrow \varepsilon$, as intended.
>
> To prove uniqueness, suppose $\varepsilon'' = (P'', [a'', b'')) \in \mathcal{E}$, R.rtperm = P'' and $\varepsilon'' \rightarrow \varepsilon$. Then P'' = R.rtperm = P'. If R.rttype = 3, then by definition, b' = a = b'', meaning both [a', b') and [a'', b'') contain $a - 1$. If R.rttype $\neq$ 3, then by definition, both [a', b') and [a'', b'') contain R.rtexdt. In both cases, Proposition 27 implies [a', b') = [a'', b''). Thus $\varepsilon' = \varepsilon''$.

If $S \subset \mathcal{E}$, then we will say that **ε is rooted in S**, abbreviated $\mathbf{S \rightarrow \varepsilon}$, if $\{\varepsilon' \in \mathcal{E} \mid \varepsilon' \rightarrow \varepsilon\} \subset S$.

That is, ε is rooted in S if all evaluators in which ε is rooted are contained in S. Note that the set $\{\varepsilon' \in \mathcal{E} \mid \varepsilon' \rightarrow \varepsilon\}$ could very well be empty (namely, whenever, and only when, $a = \min(\mathbb{M})$), in which case ε is trivially rooted in *any* subset of $\mathcal{E}$, even though ε is rooted in no particular evaluator.

Define $\mathcal{E}_0 = \{(0, \mathbb{M})\}$. For each n > 0, define $\mathcal{E}_n$ inductively by $\mathcal{E}_n = \{\varepsilon \in \mathcal{E} \mid \mathcal{E}_{n-1} \rightarrow \varepsilon\}$. Observe that $\mathcal{E}_1$ includes all evaluators rooted in no evaluators in D, in addition to all those rooted in $(0, \mathbb{M})$. That is, if $\{\varepsilon \in \mathcal{E} \mid \varepsilon \rightarrow \varepsilon'\} = \varnothing$ then $\varepsilon' \in \mathcal{E}_1$.

Proposition 30: $\mathcal{E}_n \subset \mathcal{E}_{n+1}$ for all $n \geq 0$. //Previously Propositoin 31.

> *Proof*: First observe that $\{\varepsilon \in \mathcal{E} \mid \varepsilon \rightarrow (0, \mathbb{M})\} = \varnothing \subset \{(0, \mathbb{M})\} = \mathcal{E}_0$, so $(0, \mathbb{M}) \in \mathcal{E}_1$. Thus $\mathcal{E}_0 \subset \mathcal{E}_1$.

Assume that $\mathcal{E}_k \subset \mathcal{E}_{k+1}$ for some $k \geq 0$ and fix any $\varepsilon' \in \mathcal{E}_{k+1}$. Then by definition, $\mathcal{E}_k \to \varepsilon'$, meaning $\{\varepsilon \in \mathcal{E} \mid \varepsilon \to \varepsilon'\} \subset \mathcal{E}_k$. But we are assuming $\mathcal{E}_k \subset \mathcal{E}_{k+1}$, so $\{\varepsilon \in \mathcal{E} \mid \varepsilon \to \varepsilon'\} \subset \mathcal{E}_{k+1}$. Thus $\mathcal{E}_{k+1} \to \varepsilon'$, and hence $\varepsilon' \in \mathcal{E}_{k+2}$. Since $\varepsilon' \in \mathcal{E}_{k+1}$ was arbitrary, we have proved that $\mathcal{E}_{k+1} \subset \mathcal{E}_{k+2}$. The proposition now follows by induction.

This proposition implies that $\{\mathcal{E}_n\}_{n=1}^{\infty}$ is a monotonically increasing sequence of subsets of $\mathcal{E}$. Since $\mathcal{E}$ is finite, the sequence cannot be strictly increasing. Thus there must exist some (infinitely many, in fact) $n \in \mathbb{Z}^+$ such that $\mathcal{E}_n = \mathcal{E}_{n+1}$. Let $M = \min(\{n \in \mathbb{Z}^+ \mid \mathcal{E}_n = \mathcal{E}_{n+1}\})$.

Lemma 31: If $\varepsilon = (P, [a, b))$ and $\varepsilon' = (P', [a', b'))$ are evaluators in D with $\varepsilon' \to \varepsilon$, then $a' < a$. //Previously Lemma 32.

Proof: Let R be a record in D's Roots table satisfying the conditions defining the relationship $\varepsilon' \to \varepsilon$. If R.rttype = 3 then $a = b' > a'$. If R.rttype ≠ 3 then R.rtexdt – 1 ∈ [a', b'), implying a' < R.rtexdt. By the definition of $\varepsilon' \to \varepsilon$ we have R.rtdt = a. Thus by Lemma 24.9 we have a' < R.rtexdt ≤ R.rtdt = a.

Proposition 32: $\mathcal{E} = \mathcal{E}_M$. //Previously Propositoin 33.

Proof: Assume, to obtain a contradiction, that there exists some $\varepsilon_0 \in \mathcal{E} \backslash \mathcal{E}_M$. Furthermore, suppose we have found a finite sequence $\{\varepsilon_i\} \subset \mathcal{E}$ (i = 0, 1, 2, ..., k) such that $\varepsilon_i \notin \mathcal{E}_M$ for all $0 \leq i \leq k$ and $\varepsilon_k \to \varepsilon_{k-1} \to \cdots \to \varepsilon_0$. This can certainly be done for k = 1, for otherwise if $\{\varepsilon' \in \mathcal{E} \mid \varepsilon' \to \varepsilon_0\} = \emptyset \subset \mathcal{E}_0$ then $\varepsilon_0 \in \mathcal{E}_1 \subset \mathcal{E}_M$, contradicting our assumption.

Since $\mathcal{E}_M = \mathcal{E}_{M+1}$, $\varepsilon_k \notin \mathcal{E}_{M+1}$. The latter non-containment implies that there exists some $\varepsilon_{k+1} \notin \mathcal{E}_M$ such that $\varepsilon_{k+1} \to \varepsilon_k$. Thus by induction there exists an infinite sequence $\{\varepsilon_i\} \subset \mathcal{E}$ such that $\varepsilon_i \notin \mathcal{E}_M$ for all i and $\cdots \to \varepsilon_2 \to \varepsilon_1 \to \varepsilon_0$. Lemma 31 then yields a corresponding infinite, strictly decreasing sequence of root dates in $\mathbb{M}$, which is impossible, since $\mathbb{M}$ is finite. This contradiction implies that $\mathcal{E} \backslash \mathcal{E}_M = \emptyset$, that is, that $\mathcal{E} = \mathcal{E}_M$.

Total Return Calculations: Evaluators

In this section we will define total return functions for evaluators. We will use the same notation used for accumulator total returns, relying on context for distinguishing between the two cases.

For each $\varepsilon \in \mathcal{E}$ define the function $\mathbf{Ret}_\varepsilon: \mathbb{M} \to \mathbb{R}$ as follows:

Let M be the constant referred to in Proposition 32.

For each $d \in \mathbb{M}$ where $d > \min(\mathbb{M})$, set $Ret_{(0, \mathbb{M})}(d) = |prc(0, d)|/|prc(0, d-1)|$, and set $Ret_{(0, \mathbb{M})}(\min(\mathbb{M})) = 1$. Thus $Ret_\varepsilon: \mathbb{M} \to \mathbb{R}$ is defined for all $\varepsilon \in \mathcal{E}_0$.

Assume, for some $0 < n \leq M$ and all $\varepsilon \in \mathcal{E}_{n-1}$, that $Ret_\varepsilon: \mathbb{M} \to \mathbb{R}$ is defined. Since $M = \min(\{i \geq 0 \mid \mathcal{E}_i = \mathcal{E}_{i+1}\})$, we have $\mathcal{E}_{n-1} \neq \mathcal{E}_n$, which by Proposition 30 implies $\mathcal{E}_n \backslash \mathcal{E}_{n-1} \neq \emptyset$.

Fix any $\varepsilon = (P, [a, b)) \in \mathcal{E}_n \backslash \mathcal{E}_{n-1}$. Since $\varepsilon \notin \mathcal{E}_{n-1}$ we are free to define $Ret_\varepsilon: \mathbb{M} \to \mathbb{R}$ without risk of contradiction, so fix any $x \in \mathbb{M}$.

If $x \in (b, 99991231)$ then define $Ret_\varepsilon(x) = 1$.

If $x \in (a, b]$ then by Proposition 28, $x - 1$ belongs to a tracking period for P, and thus by Proposition 16, $x - 1$ belongs to a unique (but dependent upon x) accumulation period [c, d) for P. If we let $\alpha = (P, [c, d))$, then $\alpha \in \mathcal{A}$, so we may define $Ret_\varepsilon(x) = Ret_\alpha(x)$.

If $x \in (\min(\mathbb{M}), a]$, then since a is a root date for P, the set of records in D's Roots table with permno = P and rtdt = a is nonempty by Proposition 25, so let $R_1, R_2, \ldots, R_n$ be its distinct elements. Since $a > \min(\mathbb{M})$, for each $i \in \{1, 2, ... n\}$ we may let ε_i be the unique evaluator in D corresponding to R_i by Proposition 29. The relationship $\varepsilon_i \to \varepsilon$ implies that $\varepsilon_i \in \mathcal{E}_{n-1}$ for each $i \in \{1, 2, ... n\}$ because $\varepsilon \in \mathcal{E}_n$. Thus by our induction hypothesis, $Ret_{\varepsilon i}: \mathbb{M} \to \mathbb{R}$ is defined for each $i \in \{1, 2, ... n\}$, so we may define

$$Ret_\varepsilon(x) = \frac{\sum_{i=1}^{n}\left[R_i.rtsize \div \left(\prod_{d=x+1}^{R_i.rtexdt} Ret_{\varepsilon i}(d)\right)\right]}{\sum_{i=1}^{n}\left[R_i.rtsize \div \left(\prod_{d=x}^{R_i.rtexdt} Ret_{\varepsilon i}(d)\right)\right]},$$

where

1. any product in the numerator with $R_i.rtexdt < x + 1$ is defined to be 1 and
2. any product in the denominator with $R_i.rtexdt < x$ is defined to be 1.

Finally, if $x = \min(\mathbb{M})$ then define $Ret_\varepsilon(x) = 1$.

Thus $Ret_\varepsilon: \mathbb{M} \to \mathbb{R}$ is defined for the chosen $\varepsilon \in \mathcal{E}_n \backslash \mathcal{E}_{n-1}$, and thus for all $\varepsilon \in \mathcal{E}_n \backslash \mathcal{E}_{n-1}$. By Proposition 32 and induction, $Ret_\varepsilon: \mathbb{M} \to \mathbb{R}$ is therefore defined for all $\varepsilon \in \mathcal{E}$.

Suspension and Resumption Dates

Fix any $P \in \mathcal{P}$.

Define $d \in \mathbb{M}$ to be a **suspension date** for P if
1. D prices P on d – 1 and
2. D does not price P on d.

Define $d \in \mathbb{M}$ to be a **resumption date** for P if
1. d is not a starting date for P,
2. D prices P on d, and
3. D does not price P on d – 1.

Characterization Periods

Fix any P ∈ $\mathcal{P}$.

Define d ∈ $\mathbb{M}$* to be a **critical date** for P if d is either a branch date, root date, suspension date, or resumption date for P. Note that by Proposition 14, every ending date for P is a branch date for P and thus a critical date for P, and by Proposition 25.1, every starting date for P is a root date for P and thus a critical date for P.

Define the market interval [g, h) to be a **characterization period** for P if

1. D prices P on g,
2. g and h are critical dates for P, and
3. (g, h) contains no critical dates for P.

Proposition 33: Distinct characterization periods for P are disjoint. //Previously Propositoin 34.

Proof: See the proofs of Propositions 10 and 16, which are similar.

Proposition 34: A market date d belongs to a characterization period for P if and only if D prices P on d. //Previously Propositoin 35.

Proof: Assume D prices P on d. Let G = {x ∈ (0, d] | x is a critical date for P} and let H = {x ∈ (d, 99991231] | x is a critical date for P). By Proposition 16, there exists an accumulation period [a, b) for P containing d. Since a is either a branch date or a starting date for P, a is a critical date for P, implying a ∈ G. Likewise, since b is a branch date for P, b is a critical date for P, implying b ∈ H. Since both G and H are nonempty and finite, we may define g = max(G) and h = min(H). It follows by definition that d ∈ [g, h) ⊂ [a, b), that g and h are critical dates for P, and that (g, h) = (g, d] ∪ (d, h) contains no critical dates for P.

Since D prices P on d, the sets U = {x | D prices P on y for all y ∈ [x, d]} and V = {x + 1 | D prices P on y for all y ∈ [d, x]} are nonempty and finite, so we may define u = min(U) and v = max(V).

By definition, D prices P on u but not on u – 1 (which may be 0), meaning u is either a starting date or a resumption date for P. Thus u is a critical date for P. Since u ≤ d and (g, d] contains no critical dates for P, it follows that u ≤ g, and thus that D prices P on y for all y ∈ [g, d] ⊂ [u, d].

Likewise, D prices P on v – 1 but not on v. If v = 99991231, then if follows that D prices P on y for all y ∈ [d, h). If v < 99991231, then v is a suspension date, and thus a critical date for P. Since d < v and (d, h) contains no critical dates for P, it follows that h ≤ v, and thus that D prices P on y for all y ∈ [d, h).

Thus D prices P on y for all y ∈ [g, h), and in particular, D prices P on g, making [g, h) a characterization period for P.

Conversely, assume [g, h) is a characterization period. Then by definition, D prices P on g. The proof thus far implies that g (playing the role of d) belongs to a characterization period [g', h') for P such that D prices P on y for all y ∈ [g', h'). Since g ∈ [g, h) ∩ [g', h'), Proposition 33 implies [g, h) = [g', h'). Thus D prices P on y for all y ∈ [g, h), which completes the proof.

Proposition 35: Every characterization period for P is contained within a unique accumulation period for P and a unique evaluation period for P. //Previously Propositoin 36.

> *Proof*: Let [g, h) be a characterization period for P. Since D prices P on g, by Proposition 16 there exists a unique accumulation period [a, b) for P such that $g \in [a, b)$. Likewise, by Proposition 27 there exists a unique evaluation period [c, d) for P such that $g \in [c, d)$. Since min(b, d) is a critical date for P and (g, h) contains no critical dates for P, it follows that $h \leq \min(b, d)$. Thus $[g, h) \subset [a, b)$ and $[g, h) \subset [c, d)$.

Total Return Calculations: Investments

An ordered pair $\iota = (P, [g, h))$ is defined to be an **investment** in D if $P \in \mathcal{P}$ and [g, h) is a characterization period for P. The market date g (on which ι opens for purchase) is defined to be ι's **opening date** and the market date h (on which ι closes to purchase) is defined to be ι's **closing date**. Let $\mathcal{I}$ be the set of all investments in D. Note that $(0, \mathbb{M}) \in \mathcal{I}$.

For each $\iota = (P, [g, h)) \in \mathcal{I}$ define the functions $\mathbf{Ret}_\iota\colon \mathbb{M} \to \mathbb{R}$ and $\mathbf{Liq}_\iota\colon \mathbb{M} \to \{0, 1\}$ as follows:

> Define $\mathrm{Ret}_\iota(\min(\mathbb{M})) = 1$.
>
> By Proposition 35, there exists a unique accumulation period [a, b) for P and a unique evaluation period [c, d) for P such that $[g, h) \subset [a, b) \cap [c, d)$. If we let $\alpha = (P, [a, b))$ and $\varepsilon = (P, [c, d))$, then $\alpha \in \mathcal{A}$ and $\varepsilon \in \mathcal{E}$. We may therefore define, for each $x \in (\min(\mathbb{M}), \max(\mathbb{M})]$,

$$\mathrm{Ret}_\iota(x) = \begin{cases} \mathrm{Ret}_\varepsilon(x) & \text{if } x \leq g, \\ \mathrm{Ret}_\alpha(x) & \text{if } x > g \text{ and } \mathrm{Liq}_\alpha(x-1) = \mathrm{Liq}_\alpha(x) = 1, \\ -\mathrm{Ret}_\alpha(x) & \text{if } x > g,\ \mathrm{Liq}_\alpha(x-1) = 1, \text{ and } \mathrm{Liq}_\alpha(x) = 0, \\ -\mathrm{Ret}_\alpha(x) & \text{if } x > g,\ \mathrm{Liq}_\alpha(x-1) = 0, \text{ and } \mathrm{Liq}_\alpha(x) = 1, \text{ and} \\ 1 & \text{if } x > g \text{ and } \mathrm{Liq}_\alpha(x-1) = \mathrm{Liq}_\alpha(x) = 0. \end{cases}$$

> and

$$\mathrm{Liq}_\iota(x) = \begin{cases} \mathrm{Liq}_\alpha(x) & \text{if } x \geq g \text{ and} \\ 0 & \text{if } x < g. \end{cases}$$