

Information synergy, part 1: The Kelly-Ross Theorem

In an uncertain world, synergy derives from information. For instance, information regarding interactions amongst productive inputs, outputs, or projects yields information-based economies of scale and/or scope.

A Kelly investment strategy produces maximum long-run wealth where weights, k , on Arrow-Debreu portfolios match assigned state probabilities, $k = p$, and information revising state probabilities, $\Pr(s) \equiv p$, is met with portfolio rebalancing accordingly.

Long-run wealth maximization implies maximizing the geometric mean of portfolio returns

$$\begin{aligned} \max_k \quad & \prod_{i=1}^n \left(\frac{k_i}{y_i} \right)^{\Pr(s_i)} \\ \text{s.t.} \quad & \sum_{i=1}^n k_i = 1 \end{aligned}$$

or equivalently, and perhaps more familiarly, (arithmetic) mean of logarithmic returns

$$\begin{aligned} \max_k \quad & E \left[\log \frac{k}{y} \right] \equiv \Pr(s)^T \log(\Omega k) \\ \text{s.t.} \quad & \sum_{i=1}^n k_i = 1 \end{aligned}$$

where k_i is the fraction of wealth invested in project i , $y > 0$ is a vector of no-arbitrage state prices (or Arrow-Debreu prices) derived from $Ay = x$, Ω is a diagonal matrix comprised of $\frac{1}{y_i}$, A is an $n \times n$ matrix of returns with rows referring to projects and columns to states, and x is a vector of investments (normalized to unity).¹

These notes are organized in three parts. The first part develops and briefly illustrates the combination of Ross' recovery theorem and a Kelly investment strategy culminating in the Kelly-Ross theorem. Part two explores the ex ante impact of additional coarse accounting information along with finer other information that may relate to the initial state, the post-transition state, or both. Part three focuses on ex post (probability) belief revision based on realized information signals.

Kelly criterion. The first order conditions for the Lagrangian associated with the logarithmic returns frame above regarding long-run wealth maximization is

$$\mathcal{L} = \sum_{i=1}^n p_i \ln \left(k_i \frac{1}{y_i} \right) - \lambda \left(\sum_{i=1}^n k_i - 1 \right)$$

are

$$\frac{p_i}{k_i} - \lambda = 0, \quad \text{for all } i$$

¹Weights on the nominal assets are

$$w^T = k^T \Omega A^{-1}$$

Since $\sum k_i = 1 = \sum \frac{p_i}{\lambda} = \frac{1}{\lambda}$, $\lambda = 1$ and $k_i = p_i$. In other words, probability assignment to state i identifies the optimal fractional investment in state i .

It's straightforward to show (demonstrated later) for a Kelly strategy the expected gain due to information, z , equals mutual information. Mutual information is

$$I(s; z) = H(s) + H(z) - H(s, z)$$

where $H(\cdot) = -\sum p(\cdot) \log p(\cdot)$, Shannon's entropy.²

1 Static information

Suppose returns in state i are the same regardless of the initial state. We refer to this case as a static information setting. A richer process allows returns, even riskless returns, to differ across initial states as described in Ross' recovery theorem. We regard this case as a dynamic information setting. A static information setting implies state prices and probability assignments are independent of the initial state. Consequently, the initial state can be ignored without loss of information. On the other hand, the initial state potentially carries valuable information in a dynamic information setting. This is reflected in a bounty of state-transition Arrow-Debreu (state) prices and probability assignments.

2 Dynamic information

2.1 Recovery theorem

Ross' [2011,2015]³ recovery theorem says linear no arbitrage equilibrium state prices convey a representative investor's state probability assignments. That is, state prices convey Markovian state-transition probabilities (and preferences regarding timing of consumption and risk) for a representative investor. This is in the spirit of assigning probabilities based on what we know, in other words, maximum entropy probability assignment.

The key is the pricing kernel which says the (state) price, y_{ij} , per unit probability, F_{ij} , is equal to a personal discount factor, δ , times the ratio of marginal utilities for consumption in the future state, c_j , to current, c_0 , where

²For continuous support, summation is replaced by integration, probability mass by density function, and the measure is referred to as differential entropy.

³Ross, S. 2011, "The recovery theorem," MIT working paper and Ross, S. 2015, "The recovery theorem," *The Journal of Finance*.

j refers to the future state.⁴

$$\frac{y_{ij}}{F_{ij}} = \delta \frac{U'(c_j)}{U'(c_0)}$$

In other words, a representative investor with wealth or endowment, W_0 , solves for optimal consumption subject to a budget or wealth constraint.

$$\begin{aligned} \max_{c_0, c_j \geq 0} & U(c_0) + \delta \sum_{j=1}^n F_{ij} U(c_j) \\ \text{s.t.} & c_0 + \sum_{j=1}^n y_{ij} c_j \leq W_0 \end{aligned}$$

The first order conditions for the Lagrangian representation of the above optimization problem yield the pricing kernel.

$$\begin{aligned} \lambda &= U'(c_0) \\ \delta F_{ij} U'(c_j) &= y_{ij} U'(c_0) \end{aligned}$$

For Markovian state-transition probabilities assigned as $F = \frac{1}{\delta} D P D^{-1}$ where D is a diagonal matrix with elements $U'(c_1), \dots, U'(c_n)$ and with $U'(c_0) = U'(c_i)$, then the pricing kernel for the representative investor is

$$\frac{y_{ij}}{F_{ij}} = \delta \frac{U'(c_j)}{U'(c_i)}$$

State-transition probability assignment follows from eigensystem decomposition of the dynamic system of state prices P along with the requirement the rows of F sum to one.

$$P\zeta = \delta\zeta$$

where, by the Perron-Frobenius theorem, ζ is the positive-valued eigenvector associated with the largest eigenvalue δ . The Perron-Frobenius theorem says for a nonnegative matrix the largest eigenvalue and its associated eigenvector

⁴Constant relative risk aversion is represented as

$$U(c) = \frac{c^{1-\rho}}{1-\rho}$$

where $\rho \rightarrow 1$ leads to $U(c) = \ln c$. Constant relative risk aversion is attractive as a change in wealth leads to no change in the relative composition of an individual's portfolio (fraction of wealth invested in various assets). For constant relative risk aversion, relative marginal utility is

$$\frac{U'(c_j)}{U'(c_i)} = \left(\frac{c_j}{c_i}\right)^{-\rho}$$

and logarithmic relative marginal utility is

$$\frac{U'(c_j)}{U'(c_i)} = \left(\frac{c_j}{c_i}\right)^{-1}$$

are nonnegative. Since P is a matrix of state prices, P is a positive matrix (otherwise, there exist arbitrage opportunities) with positive maximum eigenvalue and associated eigenvector.

Let ι be a vector of ones. Recall eigenvectors are scale-free, $P(\alpha\zeta) = \delta(\alpha\zeta)$ implies $P\zeta = \delta\zeta$ for any α . Then, we can write

$$D^{-1}\iota = \zeta$$

with ζ scaled appropriately. Notice, the pricing kernel is also scale-free as only ratios of marginal utilities enter. Collecting terms, we have

$$\begin{aligned} P\zeta &= \delta\zeta \\ PD^{-1}\iota &= \delta D^{-1}\iota \\ \frac{1}{\delta}DPD^{-1}\iota &= \iota \\ F\iota &= \iota \end{aligned}$$

which confirms that F is a proper probability assignment as the terms are nonnegative and sum to one.

This eigensystem decomposition of P follows directly from the pricing kernel and risk preference independence over initial states.

$$\begin{aligned} \frac{y_{ij}}{F_{ij}} &= \delta \frac{U'(c_j)}{U'(c_i)} \\ y_{ij} \frac{U'(c_i)}{U'(c_j)} &= \delta F_{ij} \end{aligned}$$

Since F_{ij} is a probability distribution given initial state i , $\sum_j F_{ij} = 1$. Therefore,

$$\begin{aligned} \sum_j y_{ij} \frac{U'(c_i)}{U'(c_j)} &= \delta \sum_j F_{ij} = \delta \\ y_{i1} \frac{U'(c_i)}{U'(c_1)} + \dots + y_{in} \frac{U'(c_i)}{U'(c_n)} &= \delta \\ y_{i1} \frac{1}{U'(c_1)} + \dots + y_{in} \frac{1}{U'(c_n)} &= \delta \frac{1}{U'(c_i)} \end{aligned}$$

For initial states, $i = 1, \dots, n$, we have n equations which in matrix form is

$$P \begin{bmatrix} \frac{1}{U'(c_1)} \\ \vdots \\ \frac{1}{U'(c_n)} \end{bmatrix} = \delta \begin{bmatrix} \frac{1}{U'(c_1)} \\ \vdots \\ \frac{1}{U'(c_n)} \end{bmatrix}$$

This is the eigensystem decomposition of P

$$PD^{-1}\iota = \delta D^{-1}\iota$$

where the eigenvector associated with the largest eigenvalue δ is

$$\begin{aligned} D^{-1}\iota &= \begin{bmatrix} \frac{1}{U'(c_1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{U'(c_n)} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{U'(c_1)} \\ \vdots \\ \frac{1}{U'(c_n)} \end{bmatrix} \end{aligned}$$

The pricing kernel, $\frac{y_{ij}}{F_{ij}} = \delta \frac{U'(c_j)}{U'(c_i)}$ or $\sum_j y_{ij} \frac{U'(c_i)}{\delta U'(c_j)} = 1$, indicates the representative agent invests $\frac{U'(c_i)}{\delta U'(c_j)}$ units in Arrow-Debreu portfolio AD_{ij} (the portfolio formed in initial state i that pays one in state j and nothing in any other state) at equilibrium unit price y_{ij} when wealth to be invested is normalized to one. In other words, the equilibrium fractional wealth invested in AD_{ij} is $y_{ij} \frac{U'(c_i)}{\delta U'(c_j)}$, but this equals the assigned probability F_{ij} by the recovery theorem and also is the long-run wealth maximizing strategy indicated by the Kelly criterion. Hence, the recovery theorem describes the equilibrium strategy for the representative investor as long-run wealth maximization.

2.2 Mutual information theorem

The mutual information theorem says the expected gain from information z equals mutual information.

$$E[r | z] - E[r] = I(s; z)$$

We consider the components and then show they are equal. First, consider expected gains from information, z , $E[r | z] - E[r]$.

$$\begin{aligned} E[r | z] &= \sum_{j=1}^n \Pr(z_j) E[r | z_j] \\ &= \sum_{j=1}^n \Pr(z_j) \sum_{i=1}^n \Pr(s_i | z_j) \log \frac{\Pr(s_i | z_j)}{y_i} \\ &= \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \log \frac{\Pr(s_i | z_j)}{y_i} \\ &= \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \log \Pr(s_i | z_j) - \sum_{i=1}^n \Pr(s_i) \log y_i \end{aligned}$$

$$\begin{aligned}
E[r] &= \sum_{i=1}^n \Pr(s_i) \log \frac{\Pr(s_i)}{y_i} \\
&= \sum_{i=1}^n \Pr(s_i) \log \Pr(s_i) - \sum_{i=1}^n \Pr(s_i) \log y_i
\end{aligned}$$

and

$$\begin{aligned}
E[r | z] - E[r] &= \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \log \Pr(s_i | z_j) - \sum_{i=1}^n \Pr(s_i) \log y_i \\
&\quad - \left(\sum_{i=1}^n \Pr(s_i) \log \Pr(s_i) - \sum_{i=1}^n \Pr(s_i) \log y_i \right) \\
&= \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \log \Pr(s_i | z_j) - \sum_{i=1}^n \Pr(s_i) \log \Pr(s_i)
\end{aligned}$$

On the other hand,

$$I(s; z) = H(s) + H(z) - H(s, z)$$

where

$$\begin{aligned}
H(s) &= - \sum_{i=1}^n \Pr(s_i) \log \Pr(s_i) \\
H(z) &= - \sum_{j=1}^n \Pr(z_j) \log \Pr(z_j)
\end{aligned}$$

and

$$\begin{aligned}
H(s, z) &= - \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \log \Pr(s_i, z_j) \\
&= - \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \{ \log \Pr(s_i | z_j) + \log \Pr(z_j) \} \\
&= - \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \log \Pr(s_i | z_j) - \sum_{j=1}^n \Pr(z_j) \log \Pr(z_j)
\end{aligned}$$

Then,

$$\begin{aligned}
I(s; z) &= - \sum_{i=1}^n \Pr(s_i) \log \Pr(s_i) - \sum_{j=1}^n \Pr(z_j) \log \Pr(z_j) \\
&\quad + \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \log \Pr(s_i | z_j) + \sum_{j=1}^n \Pr(z_j) \log \Pr(z_j) \\
&= - \sum_{i=1}^n \Pr(s_i) \log \Pr(s_i) + \sum_{i=1}^n \sum_{j=1}^n \Pr(s_i, z_j) \log \Pr(s_i | z_j)
\end{aligned}$$

Thus, we have the remarkable result — while one quantity appears denominated in returns and the other denominated in entropy, they are nevertheless equal.

$$I(s; z) = E[r | z] - E[r]$$

2.3 Scalable investments and mutual information

Construction of Arrow-Debreu portfolios from a given return profile involves scalable investments. That is, investment projects are borrowed and loaned to create "natural" (Arrow-Debreu) assets. Hence, nurturing of long-term relations with business partners facilitates such trade of productive processes. Fully exploiting the information advantage indicated by mutual information further involves rebalancing of scalable investment projects. An example illustrates these ideas.

Example 2.1 (scalable investments & mutual information) *Suppose three equally likely states and three scalable assets exhibit the following return profile.*

$$A = \begin{bmatrix} 0.5 & 1. & 1.5 \\ 1. & 1.5 & 0.5 \\ 1.5 & 0.5 & 1. \end{bmatrix} \quad x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Then, Arrow-Debreu portfolios are constructed from the rows of A^{-1} so that row one creates an Arrow-Debreu asset that pays 1 in state 1, row two creates an Arrow-Debreu asset that pays 1 in state 2, and row three creates an Arrow-Debreu asset that pays 1 in state 3.

$$A^{-1} = \begin{bmatrix} -\frac{5}{9} & \frac{1}{9} & \frac{7}{9} \\ \frac{1}{9} & \frac{7}{9} & -\frac{5}{9} \\ \frac{7}{9} & -\frac{5}{9} & \frac{1}{9} \end{bmatrix}$$

State (Arrow-Debreu) prices are

$$y = A^{-1}x = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

and expected logarithmic returns are

$$\left[\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3} \right] \log \left(\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \right) = 0$$

or expected growth equal to $\exp(0) = 1$.

Now, suppose the manager acquires perfect state information. First glance suggests the maximum growth rate is 1.5 (the maximum payoff from selecting

the asset with greatest return given the state). However, mutual information indicates the gain is $\log 3$ or, in other words, the growth rate is $\exp(0 + \log 3) = 3$. This is where scalability comes to the fore. If the information reveals state 1, then the Arrow-Debreu portfolio is 3 times row one of A^{-1} or $\left[-\frac{15}{9} \quad \frac{3}{9} \quad \frac{21}{9} \right]$ to reflect updated beliefs. Similarly, if the information reveals state 2, then the Arrow-Debreu portfolio is 3 times row two of A^{-1} or $\left[\frac{3}{9} \quad \frac{21}{9} \quad -\frac{15}{9} \right]$ and if the information reveals state 3, then the Arrow-Debreu portfolio is 3 times row three of A^{-1} or $\left[\frac{21}{9} \quad -\frac{15}{9} \quad \frac{3}{9} \right]$. Each Arrow-Debreu portfolio yields logarithmic return equal to $\log 3$ or growth rate 3 which is twice the growth rate achievable without exploiting scalability.

2.4 Kelly criterion and the recovery theorem

Suppose an investor utilizes the information in price dynamics in accordance with the recovery theorem to assign (state-transition) probabilities and employs a Kelly (long-run) investment strategy where the initial state is known prior to asset (re)allocation. The short (but remarkable) answer is the expected long-run logarithm of returns equals the logarithm of the reciprocal of the personal discount factor δ .

$$\begin{aligned} E[r | z] &= p_{ss}^T \begin{bmatrix} E[r_1] \\ E[r_2] \\ \vdots \\ E[r_n] \end{bmatrix} \\ &= p_{ss}^T \begin{bmatrix} F_1 \log(\Omega_1 F_1^T) \\ F_2 \log(\Omega_2 F_2^T) \\ \vdots \\ F_n \log(\Omega_n F_n^T) \end{bmatrix} \\ &= \log\left(\frac{1}{\delta}\right) \end{aligned}$$

where z refers to initial state information F_j is the j th row of F and $p_{ss} =$

$$\begin{bmatrix} \Pr(s_1^{ss}) \\ \Pr(s_2^{ss}) \\ \vdots \\ \Pr(s_n^{ss}) \end{bmatrix} \text{ is the long-run steady-state probability distribution such that}$$

$$p_{ss}^T F = p_{ss}^T$$

Alternatively stated, the expected growth rate in assets given a Kelly investment strategy equals the reciprocal of δ (the largest eigenvalue of the state price

matrix, P).

$$\exp(E[r | z]) = \frac{1}{\delta}$$

Hence, similar to the mutual information theorem, the Kelly-Ross combination produces an answer almost immediately to the expected long-run growth rate simply by assigning probabilities based on price information dynamics and following through with the long-run optimal investment strategy. The result is almost immediately evident when the riskless return is (initial) state-independent.

When riskless returns are equal across all initial states (state-independent), the sum of the rows of P are equal and their sum is delta. In this case, D is the identity matrix and F is a scalar multiple of P where the scalar is $\frac{1}{\delta}$. Also, the Kelly criterion generates expected growth equal to (one plus) the riskless rate in the state-independent case. Since expected returns are the same in every initial state, the long-run average expected return equals the expected return in each initial state, $\log \frac{1}{\delta}$, (steady-state probability weights play little role since it's the expected value of a constant in this state-independent case). This is illustrated in examples 2.3 and 2.4.

However, the result is quite general. It applies to the case where riskless returns are state-dependent (vary across initial states) so long as spanning is satisfied and state prices are positive (the conditions of both the recovery theorem and the Kelly criterion). This latter result is less apparent. The eigenvector of P , ζ , associated with the largest eigenvalue, δ , applies the appropriate weights to the rows of P to produce a scalar multiple of itself. Perhaps, this result is less surprising than at first blush as the recovery theorem is an equilibrium frame. If a representative agent's time preference for consumption is represented by $\frac{1}{\delta}$ then the clearing condition is that the expected growth rate equals the agent's time preference. This result is illustrated in example 2.7. Next, we consider some variations in initial state information then we state and prove the Kelly-Ross theorem.

2.4.1 Kelly criterion when initial state is ignored

As a benchmark for the value of initial state (and price dynamics) information, suppose Arrow-Debreu portfolios are formed in each initial state but the initial state information is otherwise ignored. Then, the myopic Kelly strategy is to invest wealth in the same proportions, k , as the state probabilities (i.e., the unconditional steady-state probabilities, p^{ss}). The program is

$$\begin{aligned} \max_k \quad & \sum_{j=1}^n \sum_{i=1}^n p_i^{ss} F_{ij} \log \frac{k_j}{y_{ij}} \\ \text{s.t.} \quad & \sum_{j=1}^n k_j = 1 \end{aligned}$$

After replacing k_n with $1 - k_1 - \dots - k_{n-1}$, first order conditions are

$$\begin{aligned} p_1^{ss} \left(\frac{F_{1j}}{k_j} - \frac{F_{1n}}{1 - k_1 - \dots - k_{n-1}} \right) + p_2^{ss} \left(\frac{F_{2j}}{k_j} - \frac{F_{2n}}{1 - k_1 - \dots - k_{n-1}} \right) + \\ \dots + p_n^{ss} \left(\frac{F_{nj}}{k_j} - \frac{F_{nn}}{1 - k_1 - \dots - k_{n-1}} \right) = 0 \quad j = 1, \dots, n-1 \end{aligned}$$

Solving for investment proportions k gives

$$\begin{aligned} k_j &= \frac{\sum_{i=1}^n p_i^{ss} F_{ij}}{\sum_{j=1}^n \sum_{i=1}^n p_i^{ss} F_{ij}} \\ &= \sum_{i=1}^n p_i^{ss} F_{ij} \\ &= p_j^{ss} \quad j = 1, \dots, n \end{aligned}$$

Hence, the long-run wealth maximizing strategy when the initial state is ignored is to invest fractional wealth equal to the steady-state probability, p^{ss} , in each Arrow-Debreu portfolio.

Mutual information identifies the gain from utilizing rather than ignoring initial state information.

$$\begin{aligned} E[r | z] - E[r] &= I(s, z) \\ &= H(s) + H(z) - H(s, z) \end{aligned}$$

where

$$\begin{aligned} H(s) &= -p_{ss}^T \log p_{ss} \\ H(z) &= -p_{ss}^T \log p_{ss} \\ H(s, z) &= -p(s, z)^T \log p(s, z) \end{aligned}$$

and

$$E[r | z] - E[r] = p_{ss}^T \begin{bmatrix} F_1 \log [\Omega_1 (F_1^T - p_{ss})] \\ \vdots \\ F_n \log [\Omega_n (F_n^T - p_{ss})] \end{bmatrix}$$

2.4.2 Kelly criterion when initial state is unknown

Unlike the case above where the initial state is ignored, when the initial state is unknown Arrow-Debreu portfolios cannot be identified as there are effectively n^2 states and only n assets. Let

$$A \equiv [A_1 \quad A_2 \quad \dots \quad A_n]$$

$$f \equiv \begin{bmatrix} F_1^T \\ F_2^T \\ \vdots \\ F_n^T \end{bmatrix}$$

and

$$\Omega \equiv \begin{bmatrix} \Omega_1 & 0 & 0 & 0 \\ 0 & \Omega_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \Omega_n \end{bmatrix}$$

then expected logarithmic returns when the initial state is unknown is

$$E[r_0] = \log(w^T A) \Pr(s, z)$$

s.t. $w^T \iota = \iota$

Since we can define z_0 so that $w^T A = (\Omega z_0)^T$, we can also write

$$E[r_0] = \Pr(s, z)^T \log(\Omega z_0)$$

and the expected gain due to initial state information is

$$\begin{aligned} E[r | z] - E[r_0] &= \Pr(s, z)^T \log(\Omega f) - \Pr(s, z)^T \log(\Omega z_0) \\ &= \Pr(s, z)^T [\log(\Omega f) - \log(\Omega z_0)] \end{aligned}$$

This expected gain can be identified via a "modified" mutual information.

$$\begin{aligned} E[r | z] - E[r_0] &= \hat{I}(s, z_0) \\ &\equiv H(s) + \hat{H}(z_0) - H(s, z) \end{aligned}$$

where

$$\hat{H}(z_0) = -\Pr(s, z)^T \log z_0$$

In a state-independent riskless returns setting where a riskless asset is readily identified and the initial state is unknown the long-run growth rate equals one plus the riskless rate as in the case when the initial state is known. However, when a riskless asset is formed via a portfolio of assets and this portfolio varies across initial states the long-run growth rate is less when the initial state is unknown than when it is known. This result is illustrated in example 2.5. Likewise, in a state-dependent riskless returns setting the long-run growth rate is less when the initial states are unknown than when the investor is informed. We present a corollary describing "modified" mutual information and a proof next.

Corollary 2.1 ("modified" mutual information) *The expected gain from initial state information is*

$$\begin{aligned} E[r | z] - E[r_0] &= \hat{I}(s, z_0) \\ &\equiv H(s) + \hat{H}(z_0) - H(s, z) \end{aligned}$$

where

$$\hat{H}(z_0) = -\Pr(s, z)^T \log z_0$$

Proof. First, consider the expected gain

$$\begin{aligned} E[r | z] - E[r_0] &= \Pr(s, z)^T [\log(\Omega f) - \log(\Omega z_0)] \\ &= \Pr(s, z)^T [\log \Omega + \log f - \log \Omega - \log z_0] \\ &= \Pr(s, z)^T \log f - \Pr(s, z)^T \log z_0 \end{aligned}$$

Now, "modified" mutual information is

$$\begin{aligned}\widehat{I}(s, z_0) &= H(s) + \widehat{H}(z_0) - H(s, z) \\ &= -p_{ss}^T \log p_{ss} - \Pr(s, z)^T \log z_0 + \Pr(s, z)^T \log \Pr(s, z)\end{aligned}$$

The proof is complete if we show

$$\Pr(s, z)^T \log f = -p_{ss}^T \log p_{ss} + \Pr(s, z)^T \log \Pr(s, z)$$

Since

$$\begin{aligned}\Pr(s, z)^T [\log \Pr(s, z) - \log f] &= \Pr(s, z)^T [\log p_{ss} + \log f - \log f] \\ &= p_{ss}^T \log p_{ss}\end{aligned}$$

the proof is complete.⁵ ■

Now, the stage is set for the Kelly-Ross theorem.

Theorem 2.2 (Kelly-Ross Theorem) *If state-transition probabilities are assigned in accordance with Ross' recovery theorem*

$$F = \frac{1}{\delta} D P D^{-1}$$

where spanning is satisfied, $P > 0$, and investments are made in accordance with the Kelly criterion to maximize expected long-run wealth based on known initial states (z) then the expected long-run rate of return, $E[r | z]$, equals $\log \frac{1}{\delta}$ and the expected long-run growth rate, $\exp(E[r | z])$, equals $\frac{1}{\delta}$.

Proof. Let

$$P\zeta = \delta\zeta$$

and

$$D^{-1}\iota = \zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{bmatrix}$$

⁵ A few more details regarding the last step.

$$\begin{aligned}p(s, z)^T [\log p(s, z) - \log f] &= \sum_{i=1}^n \sum_{j=1}^n p_i^{ss} F_{ij} \{\log(p_i^{ss} F_{ij}) - \log F_{ij}\} \\ &= \sum_{i=1}^n \sum_{j=1}^n p_i^{ss} F_{ij} \{\log p_i^{ss} + \log F_{ij} - \log F_{ij}\} \\ &= \sum_{i=1}^n \sum_{j=1}^n p_i^{ss} F_{ij} \log p_i^{ss} \\ &= \sum_{i=1}^n p_i^{ss} \log p_i^{ss} \\ &= p_{ss}^T \log p_{ss}\end{aligned}$$

then price dynamic information yields the state-transition probability assignment

$$F = \frac{1}{\delta} D P D^{-1}$$

where steady-state probabilities are

$$p_{ss}^T F = p_{ss}^T = [p_1^{ss} \quad p_2^{ss} \quad \cdots \quad p_n^{ss}]$$

Combining the Kelly criterion with the above probability distribution and initial state information produces

$$\begin{aligned}
E[r | z] &= p_{ss}^T \begin{bmatrix} E[r_1 | z] \\ E[r_2 | z] \\ \vdots \\ E[r_n | z] \end{bmatrix} \\
&= p_{ss}^T \begin{bmatrix} F_1 \log(\Omega_1 F_1^T) \\ F_2 \log(\Omega_2 F_2^T) \\ \vdots \\ F_n \log(\Omega_n F_n^T) \end{bmatrix} \\
&= \begin{bmatrix} p_1^{ss} F_1 \log(\Omega_1 F_1^T) \\ p_2^{ss} F_2 \log(\Omega_2 F_2^T) \\ \vdots \\ p_n^{ss} F_n \log(\Omega_n F_n^T) \end{bmatrix} \\
&= p_1^{ss} F_{11} \log \frac{1}{\delta} + p_1^{ss} F_{12} \log \frac{\zeta_2}{\delta \zeta_1} + \cdots + p_1^{ss} F_{1n} \log \frac{\zeta_n}{\delta \zeta_1} \\
&\quad + p_2^{ss} F_{21} \log \frac{\zeta_1}{\delta \zeta_2} + p_2^{ss} F_{22} \log \frac{1}{\delta} + \cdots + p_2^{ss} F_{2n} \log \frac{\zeta_n}{\delta \zeta_2} \\
&\quad + \cdots \\
&\quad + p_n^{ss} F_{n1} \log \frac{\zeta_1}{\delta \zeta_n} + p_n^{ss} F_{n2} \log \frac{\zeta_2}{\delta \zeta_n} + \cdots + p_n^{ss} F_{nn} \log \frac{1}{\delta} \\
&= \log \frac{1}{\delta} + (p_1^{ss} F_{12} - p_2^{ss} F_{21}) \log \frac{\zeta_2}{\zeta_1} + \cdots + (p_1^{ss} F_{1n} - p_n^{ss} F_{n1}) \log \frac{\zeta_n}{\zeta_1} \\
&\quad + (p_2^{ss} F_{2n} - p_n^{ss} F_{n2}) \log \frac{\zeta_n}{\zeta_2} + \cdots \\
&= \log \frac{1}{\delta} + (p_1^{ss} F_{12} - p_2^{ss} F_{21}) \log \zeta_2 + (p_2^{ss} F_{21} - p_1^{ss} F_{12}) \log \zeta_1 + \cdots \\
&\quad + (p_1^{ss} F_{1n} - p_n^{ss} F_{n1}) \log \zeta_n + (p_n^{ss} F_{n1} - p_1^{ss} F_{1n}) \log \zeta_1 \\
&\quad + (p_2^{ss} F_{2n} - p_n^{ss} F_{n2}) \log \zeta_n + (p_n^{ss} F_{n2} - p_2^{ss} F_{2n}) \log \zeta_2 + \cdots \\
&= \log \frac{1}{\delta} + (p_2^{ss} F_{21} - p_1^{ss} F_{12} + \cdots + p_n^{ss} F_{n1} - p_1^{ss} F_{1n}) \log \zeta_1 \\
&\quad + (p_1^{ss} F_{12} - p_2^{ss} F_{21} + \cdots + p_n^{ss} F_{n2} - p_2^{ss} F_{2n}) \log \zeta_2 + \cdots \\
&\quad + \left(p_1^{ss} F_{1n} - p_n^{ss} F_{n1} + p_2^{ss} F_{2n} - p_n^{ss} F_{n2} + \cdots \right) \log \zeta_n \\
&\quad \quad + p_{n-1}^{ss} F_{n-1,n} - p_n^{ss} F_{n,n-1}
\end{aligned}$$

Since

$$p_{ss}^T F = p_{ss}^T$$

then

$$\begin{aligned} p_1^{ss} F_{11} + p_2^{ss} F_{21} + \cdots + p_n^{ss} F_{n1} &= p_1^{ss} \\ p_1^{ss} F_{12} + p_2^{ss} F_{22} + \cdots + p_n^{ss} F_{n2} &= p_2^{ss} \\ &\vdots \\ p_1^{ss} F_{1n} + p_2^{ss} F_{2n} + \cdots + p_n^{ss} F_{nn} &= p_n^{ss} \end{aligned}$$

Consider the coefficients on the logarithm term, $\log \zeta_1$.

$$\begin{aligned} &(p_2^{ss} F_{21} - p_1^{ss} F_{12} + \cdots + p_n^{ss} F_{n1} - p_1^{ss} F_{1n}) \log \zeta_1 \\ &= (\{p_2^{ss} F_{21} + \cdots + p_n^{ss} F_{n1}\} - \{p_1^{ss} F_{12} + \cdots + p_1^{ss} F_{1n}\}) \log \zeta_1 \end{aligned}$$

The first term in brackets is

$$p_2^{ss} F_{21} + \cdots + p_n^{ss} F_{n1} = p_1^{ss} - p_1^{ss} F_{11}$$

while the second term in brackets is

$$p_1^{ss} F_{12} + \cdots + p_1^{ss} F_{1n} = p_1^{ss} - p_1^{ss} F_{11}$$

since F_1 is a proper probability distribution (sums to one). Hence,

$$(\{p_2^{ss} F_{21} + \cdots + p_n^{ss} F_{n1}\} - \{p_1^{ss} F_{12} + \cdots + p_1^{ss} F_{1n}\}) \log \zeta_1 = 0$$

Each of the other $\log \zeta_j$ terms is zero by analogous arguments. Therefore, the long-run expected rate of return is

$$E[r | z] = \log \frac{1}{\delta}$$

and the long-run expected growth rate is

$$\exp(E[r | z]) = \frac{1}{\delta}$$

■

The Kelly-Ross theorem indicates the maximum eigenvalue of P , δ , is the key to gauging long-run growth. Of course, this derives from the set of transition state prices and components of P , y_i . Connections amongst the components can be expressed for each initial state i as

$$\delta = \frac{1}{R_i \zeta_i} \frac{y_i \zeta}{y_i \iota}$$

where R_i is one plus the riskless return in initial state i . This result follows from two observations.

$$y_i \zeta = \delta \zeta_i$$

and

$$y_i t = \frac{1}{R_i}$$

This implies

$$R_i y_i t = \frac{y_i \zeta}{\delta \zeta_i}$$

and rearrangement yields the result above.

This result provides bounds on δ in terms of the riskless returns.

$$\min R_i \leq \frac{1}{\delta} \leq \max R_i$$

Suppose we order initial states by riskless rates such that $R_1 \leq R_2 \leq \dots \leq R_n$. Then,

$$y_i t = \frac{1}{R_i}$$

along with

$$\frac{y_i \zeta}{\zeta_i} = \delta$$

leads to $\zeta_1 \geq \zeta_2 \geq \dots \geq \zeta_n$ and

$$\frac{R_i y_i t}{\frac{y_i \zeta}{\zeta_i}} = \frac{1}{\delta}$$

Collectively, this indicates

$$R_1 \leq \frac{1}{\delta} \leq R_n$$

In a state-independent riskless returns setting, $\min R_i = \max R_i = \delta$, $\zeta_i = \zeta_j$, and $\delta = \frac{1}{R_i}$ for all i, j .

Next, we present a corollary to the Kelly-Ross theorem that says optimal long-run investment strategies when riskless returns are equal in all initial states produce zero net present value portfolios but when riskless returns vary across initial states optimal portfolios almost surely yield positive net present value.

Corollary 2.3 (net present value) *In a state-independent riskless return economy, a Kelly investment strategy results in no risk taking and a long-run growth rate equal to one plus the riskless return. In a state-dependent riskless return economy, a Kelly investment strategy results in risk taking and a long-run growth rate at least as great as one plus the riskless return in each initial state.*

First, since spanning is satisfied (along with positive state prices or no arbitrage) the riskless portfolio is always achievable when the initial state is known and the Kelly strategy is optimal in the long-run, a Kelly investor cannot do worse in expectation than the riskless return. The pricing kernel indicates that an investor cannot do better than the riskless return in a state-independent riskless return economy as the riskless return is the same in all initial states.

However, the pricing kernel reveals there are potential gains in a state-dependent riskless return economy.

Proof. The state-independent riskless returns case is immediate from the Kelly-Ross theorem. A Kelly strategy generates growth equal to $\frac{1}{\delta}$ in each initial state and one plus the riskless rate of return, R , equals $\frac{1}{\delta}$ in all states. The expected logarithmic return in initial state i given initial state knowledge is

$$E[r_i | z] = F_{i1} \log \frac{\zeta_1}{\delta \zeta_i} + \cdots + F_{in} \log \frac{\zeta_n}{\delta \zeta_i}$$

Since $\zeta_i = \zeta_j$ for all i, j ,

$$\begin{aligned} E[r_i | z] &= F_{i1} \log \frac{1}{\delta} + \cdots + F_{in} \log \frac{1}{\delta} \\ &= \log \frac{1}{\delta} \\ &= \log(1 + r^f) \end{aligned}$$

The state-dependent riskless returns case also follows from the theorem but in this informationally richer setting the pricing kernel allows greater returns in all except knife-edge cases.

$$\begin{aligned} E[r_i | z] &= F_{i1} \log \frac{\zeta_1}{\delta \zeta_i} + \cdots + F_{in} \log \frac{\zeta_n}{\delta \zeta_i} \\ &= \frac{\zeta_1 y_{i1}}{\delta \zeta_i} \log \frac{\zeta_1}{\delta \zeta_i} + \cdots + \frac{\zeta_n y_{in}}{\delta \zeta_i} \log \frac{\zeta_n}{\delta \zeta_i} \end{aligned}$$

where

$$\sum_{j=1}^n y_{ij} = \frac{1}{1 + r_i^f}$$

and

$$y_i \zeta = \delta \zeta \quad \text{for all } i$$

Since the Kelly criterion is optimal given the initial state and spanning ensures the riskless return is accessible, a Kelly investor generates an expected growth rate at least as large as one plus the riskless rate. ■

To illustrate the corollary for the state-dependent riskless returns case, consider a knife-edge (pathological) example. Suppose

$$A_1 = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ -1 & 1.5 \end{bmatrix}$$

and

$$A_2 = \begin{bmatrix} 1.1 & 1.1 \\ 1.1 & \frac{1}{1.1} \end{bmatrix}$$

Then,

$$P = \begin{bmatrix} 0.5 & 1 \\ \frac{1}{1.1} & 0 \end{bmatrix}$$

and

$$F = \begin{bmatrix} 0.404631 & 0.595369 \\ 1 & 0 \end{bmatrix}$$

This produces expected logarithmic returns and growth rates as follows.

$$E[r_1 | z] = -0.394376$$

$$\exp(E[r_1 | z]) = 0.674101 > \frac{2}{3}$$

and

$$E[r_2 | z] = 0.0953102$$

$$\exp(E[r_2 | z]) = 1.1$$

Hence, the growth in initial state two equals the riskless return but the growth rate for initial state one is greater than the riskless return. Knife-edge cases such as this one exist but, as subsequent examples illustrate, the long-run growth rate typically exceeds one plus the riskless rate in all initial states for the state-dependent riskless return setting.

Now, we illustrate the combination of Ross' recovery theorem probability assignment with a Kelly (long-term wealth maximizing) investment strategy with some numerical examples.

Example 2.2 (static information) *Suppose there are three assets and three states*

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1.1 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.1 \end{bmatrix}, x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

and $y = \begin{bmatrix} 0.332326 \\ 0.365559 \\ 0.302115 \end{bmatrix}$. *As returns are static (don't depend on the initial state), the state price matrix involves the same row repeated*

$$P = \begin{bmatrix} 0.332326 & 0.365559 & 0.302115 \\ 0.332326 & 0.365559 & 0.302115 \\ 0.332326 & 0.365559 & 0.302115 \end{bmatrix}$$

and since the riskless rate is zero, state-transition probabilities are equal to P .

$$\begin{aligned}
F &= \frac{1}{\delta} D P D^{-1} \\
&= \frac{1}{1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.332326 & 0.365559 & 0.302115 \\ 0.332326 & 0.365559 & 0.302115 \\ 0.332326 & 0.365559 & 0.302115 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 0.332326 & 0.365559 & 0.302115 \\ 0.332326 & 0.365559 & 0.302115 \\ 0.332326 & 0.365559 & 0.302115 \end{bmatrix}
\end{aligned}$$

Since $\delta = 1$, $E[r | z] = \log \frac{1}{1} = 0$ and the expected long-run growth rate is $\exp(0) = \frac{1}{\delta} = 1$ as indicated by the Kelly-Ross theorem.

Next, we illustrate a dynamic information setting but with state-independent riskless returns. That is, riskless returns are the same in all initial states.

Example 2.3 (dynamic information, state-independent riskless returns)

Suppose returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1.03 & 1.03 & 1.03 \\ 1.1 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.1 \end{bmatrix}$$

initial state two returns are

$$A_2 = \begin{bmatrix} 1.03 & 1.03 & 1.03 \\ 1.2 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.2 \end{bmatrix}$$

and for initial state three returns are

$$A_3 = \begin{bmatrix} 1.03 & 1.03 & 1.03 \\ 1.3 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.3 \end{bmatrix}$$

Then, the matrix of state-transition prices is

$$P = \begin{bmatrix} 0.332326 & 0.0451706 & 0.593377 \\ 0.313481 & 0.36927 & 0.288122 \\ 0.259432 & 0.535738 & 0.175703 \end{bmatrix}$$

where

$$P \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0.970874 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

and the matrix of state-transition probabilities is

$$\begin{aligned}
F &= \frac{1}{0.970874} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&\times \begin{bmatrix} 0.332326 & 0.0451706 & 0.593377 \\ 0.313481 & 0.36927 & 0.288122 \\ 0.259432 & 0.535738 & 0.175703 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0.342296 & 0.0465257 & 0.611178 \\ 0.322886 & 0.380348 & 0.296766 \\ 0.267215 & 0.55181 & 0.180974 \end{bmatrix}
\end{aligned}$$

Steady-state probabilities (initial state probabilities equal state probabilities following transition given state-transition probabilities, F)⁶ are

$$\begin{aligned}
p_{ss}^T F &= p_{ss}^T \\
p_{ss}^T &= [0.30923 \quad 0.337664 \quad 0.353105]
\end{aligned}$$

A Kelly investment strategy produces an expected return in initial state one equal to

$$\begin{aligned}
E[r_1 | z] &= [0.342296 \quad 0.0465257 \quad 0.611178] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.0451706} & 0 \\ 0 & 0 & \frac{1}{0.593377} \end{bmatrix} \begin{bmatrix} 0.342296 \\ 0.0465257 \\ 0.611178 \end{bmatrix} \right) \\
&= 0.0295588
\end{aligned}$$

or expected periodic growth in value equal to $\exp[0.0295588] = 1.03$, initial state two involves

$$\begin{aligned}
E[r_2 | z] &= [0.322886 \quad 0.380348 \quad 0.296766] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.313481} & 0 & 0 \\ 0 & \frac{1}{0.36927} & 0 \\ 0 & 0 & \frac{1}{0.288122} \end{bmatrix} \begin{bmatrix} 0.322886 \\ 0.380348 \\ 0.296766 \end{bmatrix} \right) \\
&= 0.0295588
\end{aligned}$$

or expected periodic growth in value equal to $\exp[0.0295588] = 1.03$, and initial state three involves

$$\begin{aligned}
E[r_3 | z] &= [0.267215 \quad 0.55181 \quad 0.180974] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.259432} & 0 & 0 \\ 0 & \frac{1}{0.535738} & 0 \\ 0 & 0 & \frac{1}{0.175703} \end{bmatrix} \begin{bmatrix} 0.267215 \\ 0.55181 \\ 0.180974 \end{bmatrix} \right) \\
&= 0.0295588
\end{aligned}$$

⁶Steady-state probabilities are the normalized (sum to one) eigenvector associated with a unit eigenvalue for F^T , $F^T p_{ss} = 1 p_{ss}$. In this setting, the state probability distribution converges quickly.

or expected periodic growth in value equal to $\exp[0.0295588] = 1.03$. Hence, expected steady-state return is

$$\begin{aligned} E[r | z] &= p_{ss}^T \begin{bmatrix} E[r_1 | z] \\ E[r_2 | z] \\ E[r_3 | z] \end{bmatrix} \\ &= \begin{bmatrix} 0.30923 & 0.337664 & 0.353105 \end{bmatrix} \begin{bmatrix} 0.0295588 \\ 0.0295588 \\ 0.0295588 \end{bmatrix} \\ &= 0.0295588 \end{aligned}$$

or expected periodic growth $\exp[0.0295588] = 1.03 = \frac{1}{\delta} = \frac{1}{0.970874}$ as indicated by the Kelly-Ross theorem. Briefly, mutual information indicates initial state information is quite valuable as $I(s; z) = 0.127712$ or growth equal to $\exp(0.127712) = 1.13623$.

Spanning supplies a derived riskless asset even when it's not obvious one exists. If the derived riskless rates are the same across all initial states then the long-run expected growth rate equals one plus this riskless rate or $\frac{1}{\delta}$ as in the foregoing example. This case is illustrated next with minor modification of example 2.3.

Example 2.4 (derived state-independent riskless return) Suppose returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1.04 & 1.03 & 1.0244 \\ 1.1 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.1 \end{bmatrix}$$

initial state two returns are

$$A_2 = \begin{bmatrix} 1.04 & 1.03 & 1.01912 \\ 1.2 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.2 \end{bmatrix}$$

and for initial state three returns are

$$A_3 = \begin{bmatrix} 1.04 & 1.03 & 1.01523 \\ 1.3 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.3 \end{bmatrix}$$

Then, the matrix of state-transition prices is

$$P = \begin{bmatrix} 0.332326 & 0.0451706 & 0.593377 \\ 0.313481 & 0.36927 & 0.288122 \\ 0.259432 & 0.535738 & 0.175703 \end{bmatrix}$$

where

$$P \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0.970874 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

and the matrix of state-transition probabilities is

$$\begin{aligned}
F &= \frac{1}{0.970874} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&\times \begin{bmatrix} 0.332326 & 0.0451706 & 0.593377 \\ 0.313481 & 0.36927 & 0.288122 \\ 0.259432 & 0.535738 & 0.175703 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0.342296 & 0.0465257 & 0.611178 \\ 0.322886 & 0.380348 & 0.296766 \\ 0.267215 & 0.55181 & 0.180974 \end{bmatrix}
\end{aligned}$$

Steady-state probabilities are

$$\begin{aligned}
p_{ss}^T F &= p_{ss}^T \\
p_{ss}^T &= [0.30923 \quad 0.337664 \quad 0.353105]
\end{aligned}$$

The derived riskless portfolio for initial state one is also the long-run optimal investment (Kelly) strategy.

$$\begin{aligned}
w_1^T &= F_1 \Omega_1 A_1^{-1} \\
&= [0.948428 \quad -0.0170178 \quad 0.0685903]
\end{aligned}$$

with constant (riskless) state-by-state returns

$$w_1^T A_1 = [1.03 \quad 1.03 \quad 1.03]$$

The derived riskless portfolio for initial state two is

$$\begin{aligned}
w_2^T &= F_2 \Omega_2 A_2^{-1} \\
&= [0.955926 \quad -0.0145426 \quad 0.0586164]
\end{aligned}$$

with constant (riskless) state-by-state returns

$$w_2^T A_2 = [1.03 \quad 1.03 \quad 1.03]$$

The derived riskless portfolio for initial state three is

$$\begin{aligned}
w_3^T &= F_3 \Omega_3 A_3^{-1} \\
&= [0.961506 \quad -0.0126929 \quad 0.051187]
\end{aligned}$$

with constant (riskless) state-by-state returns

$$w_3^T A_3 = [1.03 \quad 1.03 \quad 1.03]$$

A Kelly investment strategy produces an expected return in initial state one equal to

$$\begin{aligned}
E[r_1 | z] &= [0.342296 \quad 0.0465257 \quad 0.611178] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.0451706} & 0 \\ 0 & 0 & \frac{1}{0.593377} \end{bmatrix} \begin{bmatrix} 0.342296 \\ 0.0465257 \\ 0.611178 \end{bmatrix} \right) \\
&= 0.0295588
\end{aligned}$$

or expected periodic growth in value equal to $\exp [0.0295588] = 1.03$, initial state two involves

$$\begin{aligned} E[r_2 | z] &= [0.322886 \quad 0.380348 \quad 0.296766] \\ &\times \log \left(\left[\begin{array}{ccc} \frac{1}{0.313481} & 0 & 0 \\ 0 & \frac{1}{0.36927} & 0 \\ 0 & 0 & \frac{1}{0.288122} \end{array} \right] \left[\begin{array}{c} 0.322886 \\ 0.380348 \\ 0.296766 \end{array} \right] \right) \\ &= 0.0295588 \end{aligned}$$

or expected periodic growth in value equal to $\exp [0.0295588] = 1.03$, and initial state three involves

$$\begin{aligned} E[r_3 | z] &= [0.267215 \quad 0.55181 \quad 0.180974] \\ &\times \log \left(\left[\begin{array}{ccc} \frac{1}{0.259432} & 0 & 0 \\ 0 & \frac{1}{0.535738} & 0 \\ 0 & 0 & \frac{1}{0.175703} \end{array} \right] \left[\begin{array}{c} 0.267215 \\ 0.55181 \\ 0.180974 \end{array} \right] \right) \\ &= 0.0295588 \end{aligned}$$

or expected periodic growth in value equal to $\exp [0.0295588] = 1.03$. Hence, expected steady-state return is

$$\begin{aligned} E[r | z] &= p_{ss}^T \begin{bmatrix} E[r_1 | z] \\ E[r_2 | z] \\ E[r_3 | z] \end{bmatrix} \\ &= [0.30923 \quad 0.337664 \quad 0.353105] \begin{bmatrix} 0.0295588 \\ 0.0295588 \\ 0.0295588 \end{bmatrix} \\ &= 0.0295588 \end{aligned}$$

or expected periodic growth $\exp [0.0295588] = 1.03 = \frac{1}{\delta} = \frac{1}{0.970874}$ as indicated by the Kelly-Ross theorem.

Example 2.5 ("modified" mutual information) Continue with example 2.4. Now, suppose the initial state is unknown. Clearly, an investor cannot replicate the above Kelly strategy as different portfolios across initial states produce riskless returns. However, long-run expected logarithmic returns are maximized via

$$\begin{aligned} \max_w \quad & \log(w^T A) \Pr(s, z) \\ \text{s.t.} \quad & w^T \iota = 1 \end{aligned}$$

The solution is

$$w = \begin{bmatrix} 0.95591 \\ -0.01334 \\ 0.05743 \end{bmatrix}$$

with

$$w^T A =$$

[1.0317 1.0299 1.0291 1.0304 1.0299 1.0298 1.0290 1.0299 1.0318]

or little variation in returns (but nonconstant returns). Now, solve

$$w^T A = (\Omega z_0)^T$$

$$\begin{bmatrix} w_1 & w_2 & 1 - w_1 - w_2 \end{bmatrix} \begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix} = \left(\begin{bmatrix} \Omega_1 & 0 & 0 \\ 0 & \Omega_2 & 0 \\ 0 & 0 & \Omega_3 \end{bmatrix} z_0 \right)^T$$

for z_0 . This gives $z_0^T =$

[0.3429 0.0465 0.6106 0.3230 0.3803 0.2967 0.2670 0.5518 0.1813]

Hence,

$$\begin{aligned} E[r_0] &= \Pr(s, z)^T \log(\Omega z_0) \\ &= 0.0295584 \end{aligned}$$

and the expected gain from initial state information is

$$\begin{aligned} E[r | z] - E[r_0] &= \log 1.03 - 0.0295584 \\ &= 0.0295588 - 0.0295584 \\ &= 0.0000004 \end{aligned}$$

which equals "modified" mutual information

$$\begin{aligned} \hat{I}(s, z_0) &= H(s) + \hat{H}(z_0) - H(s, z) \\ &= 1.09712 + 0.969404 - 2.06652 \\ &= 0.0000004 \end{aligned}$$

or long-run expected growth differential, $\exp(0.0000004) = 1.0000004$.

State-dependent riskless returns present a greater challenge. However, the expected long-run periodic growth rate continues to be equal to $\frac{1}{5}$ as indicated by the Kelly-Ross theorem. This case is illustrated next along with discussion of the value of initial state information.

Example 2.6 (mutual information and state-dependent riskless returns)

Suppose returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1.01 & 1.01 \\ 1.1 & \frac{1}{1.1} \end{bmatrix}$$

and for initial state two returns are

$$A_2 = \begin{bmatrix} 1.05 & 1.05 \\ \frac{1}{1.1} & 1.1 \end{bmatrix}$$

Then, the matrix of state-transition prices is

$$P = \begin{bmatrix} 0.523338 & 0.466761 \\ 0.249433 & 0.702948 \end{bmatrix}$$

and the matrix of state-transition probabilities is

$$\begin{aligned} F &= \frac{1}{0.965975} \begin{bmatrix} 1.37815 & 0 \\ 0 & 1.45326 \end{bmatrix} \\ &\times \begin{bmatrix} 0.523338 & 0.466761 \\ 0.249433 & 0.702948 \end{bmatrix} \begin{bmatrix} 0.725609 & 0 \\ 0 & 0.688107 \end{bmatrix} \\ &= \begin{bmatrix} 0.541772 & 0.458228 \\ 0.272292 & 0.727708 \end{bmatrix} \end{aligned}$$

Steady-state probabilities are

$$\begin{aligned} p_{ss}^T F &= p_{ss}^T \\ p_{ss}^T &= [0.372737 \quad 0.627263] \end{aligned}$$

The long-run optimal investment (Kelly) strategy for initial state one is

$$\begin{aligned} w_1^T &= F_1 \Omega_1 A_1^{-1} \\ &= [0.719741 \quad 0.280259] \end{aligned}$$

with nonconstant (unlike the state-independent case) state-by-state returns

$$w_1^T A_1 = [1.03522 \quad 0.981719]$$

The long-run optimal investment (Kelly) strategy for initial state two is

$$\begin{aligned} w_2^T &= F_2 \Omega_2 A_2^{-1} \\ &= [1.29553 \quad -0.295533] \end{aligned}$$

with nonconstant state-by-state returns

$$w_2^T A_2 = [1.09164 \quad 1.03522]$$

A Kelly investment strategy produces an expected return in initial state one equal to

$$\begin{aligned} E[r_1 | z] &= [0.541772 \quad 0.458228] \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.523338} & 0 \\ 0 & \frac{1}{0.466761} \end{bmatrix} \begin{bmatrix} 0.541772 \\ 0.458228 \end{bmatrix} \right) \\ &= 0.0103004 \end{aligned}$$

or expected periodic growth in value equal to $\exp[0.0103004] = 1.01035$, while initial state two involves

$$\begin{aligned} E[r_2 | z] &= [0.272292 \quad 0.727708] \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.249433} & 0 \\ 0 & \frac{1}{0.702948} \end{bmatrix} \begin{bmatrix} 0.272292 \\ 0.727708 \end{bmatrix} \right) \\ &= 0.0490669 \end{aligned}$$

or expected periodic growth in value equal to $\exp [0.0490669] = 1.05029$. Hence, expected steady-state return is

$$\begin{aligned} E[r | z] &= p_{ss}^T \begin{bmatrix} E[r_1] \\ E[r_2] \end{bmatrix} \\ &= \begin{bmatrix} 0.372737 & 0.627263 \end{bmatrix} \begin{bmatrix} 0.0103004 \\ 0.0490669 \end{bmatrix} \\ &= 0.0346172 \end{aligned}$$

or expected periodic growth is $\exp [0.0346172] = 1.03522 = \frac{1}{\delta} = \frac{1}{0.965975}$ as indicated by the Kelly-Ross theorem. The expected value of state-dependent initial state information is reflected in mutual information. The long-run joint distribution of transition and initial states assigned to reflect the price dynamics of this setting is

$$\begin{aligned} &\begin{bmatrix} 0.372737 \times 0.541772 & 0.372737 \times 0.458228 \\ 0.627263 \times 0.272292 & 0.627263 \times 0.727708 \end{bmatrix} \\ &= \begin{bmatrix} 0.201938 & 0.170799 \\ 0.170799 & 0.456464 \end{bmatrix} \end{aligned}$$

where rows refer to initial state information (z) and columns refer to states (s). Thus, mutual information is

$$\begin{aligned} I(s; z) &= H(s) + H(z) - H(s, z) \\ &= 0.660396 + 0.660396 - 1.28473 \\ &= 0.0360587 \end{aligned}$$

or expected gains in the growth rate due to the initial state information in price dynamics equals $\exp [0.0360587] = 1.03672$. This quantity is confirmed by comparing expected returns based on a Kelly strategy employing initial state information ($E[r | z] = \log \frac{1}{\delta} = 0.0346172$) with a Kelly strategy that ignores the initial state information. The latter strategy involves a constant investment strategy (ignoring initial state information) reflecting steady-state probabilities that generates an expected return in initial state one equal to

$$\begin{aligned} E[r_1] &= \begin{bmatrix} 0.541772 & 0.458228 \end{bmatrix} \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.523338} & 0 \\ 0 & \frac{1}{0.466761} \end{bmatrix} \begin{bmatrix} 0.372737 \\ 0.627263 \end{bmatrix} \right) \\ &= -0.0484239 \end{aligned}$$

or expected periodic growth in value equal to $\exp [-0.0484239] = 0.95273$, and initial state two generates

$$\begin{aligned} E[r_2] &= \begin{bmatrix} 0.272292 & 0.727708 \end{bmatrix} \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.249433} & 0 \\ 0 & \frac{1}{0.702948} \end{bmatrix} \begin{bmatrix} 0.372737 \\ 0.627263 \end{bmatrix} \right) \\ &= 0.0264767 \end{aligned}$$

or expected periodic growth in value equal to $\exp [0.0264767] = 1.02683$. Hence, expected steady-state returns ignoring the initial state information is

$$\begin{aligned} E[r] &= p_{ss}^T \begin{bmatrix} E[r_1] \\ E[r_2] \end{bmatrix} \\ &= \begin{bmatrix} 0.372737 & 0.627263 \end{bmatrix} \begin{bmatrix} -0.0484239 \\ 0.0264767 \end{bmatrix} \\ &= -0.00144155 \end{aligned}$$

or expected periodic growth $\exp [-0.00144155] = 0.998559$. As indicated by the mutual information theorem, expected gains from the initial state information equals mutual information.

$$\begin{aligned} E[r | z] - E[r] &= 0.0346172 - (-0.00144155) \\ &= 0.0360587 \\ &= I(s; z) \end{aligned}$$

Example 2.7 (mutual information — expanded example) Suppose returns associated with initial state one are

$$A_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1.1 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.1 \end{bmatrix}$$

initial state two returns are

$$A_2 = \begin{bmatrix} 1.05 & 1.05 & 1.05 \\ 1.2 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.2 \end{bmatrix}$$

and for initial state three returns are

$$A_3 = \begin{bmatrix} 1.1 & 1.1 & 1.1 \\ 1.3 & \frac{1}{1.1} & 1 \\ \frac{1}{1.1} & 1 & 1.3 \end{bmatrix}$$

Then, the matrix of state-transition prices is

$$P = \begin{bmatrix} 0.332326 & 0.365559 & 0.302115 \\ 0.338782 & 0.221511 & 0.392087 \\ 0.348914 & 0.151415 & 0.408762 \end{bmatrix}$$

and the matrix of state-transition probabilities is

$$\begin{aligned}
F &= \frac{1}{0.952747} \begin{bmatrix} 1.64931 & 0 & 0 \\ 0 & 1.73825 & 0 \\ 0 & 0 & 1.82142 \end{bmatrix} \\
&\times \begin{bmatrix} 0.332326 & 0.365559 & 0.302115 \\ 0.338782 & 0.221511 & 0.392087 \\ 0.348914 & 0.151415 & 0.408762 \end{bmatrix} \begin{bmatrix} 0.606314 & 0 & 0 \\ 0 & 0.57529 & 0 \\ 0 & 0 & 0.549021 \end{bmatrix} \\
&= \begin{bmatrix} 0.348809 & 0.364057 & 0.287135 \\ 0.374761 & 0.232498 & 0.392742 \\ 0.404436 & 0.166529 & 0.429035 \end{bmatrix}
\end{aligned}$$

Steady-state probabilities are

$$\begin{aligned}
p_{ss}^T F &= p_{ss}^T \\
p_{ss}^T &= [0.375877 \quad 0.257781 \quad 0.366342]
\end{aligned}$$

The long-run optimal investment (Kelly) strategy for initial state one is

$$\begin{aligned}
w_1^T &= F_1 \Omega_1 A_1^{-1} \\
&= [1.45064 \quad 0.0452009 \quad -0.495843]
\end{aligned}$$

with nonconstant state-by-state returns

$$w_1^T A_1 = [1.0496 \quad 0.995891 \quad 0.950416]$$

The Kelly strategy for initial state two is

$$\begin{aligned}
w_2^T &= F_2 \Omega_2 A_2^{-1} \\
&= [1.18248 \quad 0.104797 \quad -0.287273]
\end{aligned}$$

with nonconstant state-by-state returns

$$w_2^T A_2 = [1.1062 \quad 1.0496 \quad 1.00167]$$

The Kelly strategy for initial state three is

$$\begin{aligned}
w_3^T &= F_3 \Omega_3 A_3^{-1} \\
&= [1.09396 \quad 0.105368 \quad -0.199332]
\end{aligned}$$

with nonconstant state-by-state returns

$$w_3^T A_3 = [1.15913 \quad 1.09982 \quad 1.0496]$$

A Kelly investment strategy produces an expected return in initial state one equal to

$$\begin{aligned}
E[r_1 | z] &= [0.348809 \quad 0.364057 \quad 0.287135] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.365559} & 0 \\ 0 & 0 & \frac{1}{0.302115} \end{bmatrix} \begin{bmatrix} 0.348809 \\ 0.364057 \\ 0.287135 \end{bmatrix} \right) \\
&= 0.000782916
\end{aligned}$$

or expected periodic growth in value equal to $\exp [0.000782916] = 1.00078$, initial state two involves

$$\begin{aligned} E[r_2 | z] &= [0.374761 \quad 0.232498 \quad 0.392742] \\ &\times \log \left(\left[\begin{array}{ccc} \frac{1}{0.338782} & 0 & 0 \\ 0 & \frac{1}{0.221511} & 0 \\ 0 & 0 & \frac{1}{0.392087} \end{array} \right] \left[\begin{array}{c} 0.374761 \\ 0.232498 \\ 0.392742 \end{array} \right] \right) \\ &= 0.0497338 \end{aligned}$$

or expected periodic growth in value equal to $\exp [0.0497338] = 1.05099$, and initial state three involves

$$\begin{aligned} E[r_3 | z] &= [0.404436 \quad 0.166529 \quad 0.429035] \\ &\times \log \left(\left[\begin{array}{ccc} \frac{1}{0.348914} & 0 & 0 \\ 0 & \frac{1}{0.151415} & 0 \\ 0 & 0 & \frac{1}{0.408762} \end{array} \right] \left[\begin{array}{c} 0.404436 \\ 0.166529 \\ 0.429035 \end{array} \right] \right) \\ &= 0.0963344 \end{aligned}$$

or expected periodic growth in value equal to $\exp [0.0963344] = 1.10113$. Hence, expected steady-state return is

$$\begin{aligned} E[r | z] &= p_{ss}^T \begin{bmatrix} E[r_1 | z] \\ E[r_2 | z] \\ E[r_3 | z] \end{bmatrix} \\ &= [0.375877 \quad 0.257781 \quad 0.366342] \begin{bmatrix} 0.000782916 \\ 0.0497338 \\ 0.0963344 \end{bmatrix} \\ &= 0.048406 \end{aligned}$$

or expected periodic growth $\exp [0.048406] = 1.0496 = \frac{1}{\delta} = \frac{1}{0.952747}$ as indicated by the Kelly-Ross theorem. Let's explore this long-run expected growth rate from another perspective. Suppose state realizations arise at the steady-state frequency. Four draws produce

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

where rows refer to initial states and columns refer to state transitions. An initial investment of 1 grows to 1.27171 or a geometric mean (average periodic growth rate) equal to $(1.27171)^{\frac{1}{4}} = 1.06193$. Seven draws produce

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

An initial investment of 1 grows to 1.33924 or a geometric mean (average periodic growth rate) equal to $(1.33924)^{\frac{1}{7}} = 1.04261$. Ten draws produce

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

An initial investment of 1 grows to 1.62265 or a geometric mean (average periodic growth rate) equal to $(1.62265)^{\frac{1}{10}} = 1.0496 \approx \frac{1}{\delta}$. Hence, this setting converges quickly to the expected long-run growth rate. If initial state information is ignored a Kelly investment strategy invests fractions of wealth equal to the steady-state probabilities resulting in the following expected returns by initial state.

$$\begin{aligned} E[r_1] &= [0.348809 \quad 0.364057 \quad 0.287135] \\ &\quad \times \log \left(\begin{bmatrix} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.365559} & 0 \\ 0 & 0 & \frac{1}{0.302115} \end{bmatrix} \begin{bmatrix} 0.375877 \\ 0.257781 \\ 0.366342 \end{bmatrix} \right) \\ &= -0.0288693 \end{aligned}$$

$$\begin{aligned} E[r_2] &= [0.374761 \quad 0.232498 \quad 0.392742] \\ &\quad \times \log \left(\begin{bmatrix} \frac{1}{0.338782} & 0 & 0 \\ 0 & \frac{1}{0.221511} & 0 \\ 0 & 0 & \frac{1}{0.392087} \end{bmatrix} \begin{bmatrix} 0.375877 \\ 0.257781 \\ 0.366342 \end{bmatrix} \right) \\ &= 0.0475205 \end{aligned}$$

$$\begin{aligned} E[r_3] &= [0.404436 \quad 0.166529 \quad 0.429035] \\ &\quad \times \log \left(\begin{bmatrix} \frac{1}{0.348914} & 0 & 0 \\ 0 & \frac{1}{0.151415} & 0 \\ 0 & 0 & \frac{1}{0.408762} \end{bmatrix} \begin{bmatrix} 0.375877 \\ 0.257781 \\ 0.366342 \end{bmatrix} \right) \\ &= 0.0717054 \end{aligned}$$

and

$$\begin{aligned} E[r] &= [E[r_1] \quad E[r_2] \quad E[r_3]] p_{ss} \\ &= [-0.0288693 \quad 0.0475205 \quad 0.0717054] \begin{bmatrix} 0.375877 \\ 0.257781 \\ 0.366342 \end{bmatrix} \\ &= 0.0276672 \end{aligned}$$

Hence, the expected gain from utilizing the information in initial states is $E[r | z] - E[r] = 0.048406 - 0.0276672 = 0.0207388$. Of course, this equals mutual information. The joint distribution of state (columns) and information (rows) is

$$\Pr(s, z) = \begin{bmatrix} 0.131109 & 0.136841 & 0.107927 \\ 0.0966061 & 0.0599334 & 0.101241 \\ 0.148162 & 0.0610067 & 0.157174 \end{bmatrix}$$

The probability distribution over the information (initial state, z) equals the sum of the rows which produces the steady-state probabilities while the probability distribution over the states (s) is the sum of the columns which also equals the steady-state probabilities. Consequently, mutual information is

$$\begin{aligned} I(s; z) &= H(s) + H(z) - H(s, z) \\ &= 1.08513 + 1.08513 - 2.14952 \\ &= 0.0207388 \end{aligned}$$

equal to expected gains as indicated by the mutual information theorem, or expected gains in long-run growth rate due to initial state information is equal to

$$\exp(0.0207388) = 1.02096$$

Next, we consider the expected marginal (or incremental) gain from additional information.⁷

2.5 Mutual marginal information

Suppose z_1 information is in place (such as initial state information), what is the expected marginal or incremental gain from additional information z_2 ? Unlike one-shot opportunities, long-run wealth maximization makes this straightforward via a direct extension of the mutual information theorem. The expected marginal gain is

$$\begin{aligned} E[\text{gain}(z_2 | z_1)] &= I(s; z_2 | z_1) \\ &= H(s | z_1) + H(z_2 | z_1) - H(s, z_2 | z_1) \\ &= I(s; z_1, z_2) - I(s; z_1) \end{aligned}$$

The connection between the second and third lines follows primarily from entropy additivity: $H(x, y) = H(x | y) + H(y) = H(y | x) + H(x)$.

$$\begin{aligned} I(s; z_1, z_2) &= H(s) + H(z_1, z_2) - H(s, z_1, z_2) \\ &= H(s) + H(z_2 | z_1) + H(z_1) - H(s, z_2 | z_1) - H(z_1) \\ &= H(s) + H(z_2 | z_1) - H(s, z_2 | z_1) \end{aligned}$$

⁷Part 2 reports more extensive examples of accounting and other initial state and/or post-transition state information.

and

$$\begin{aligned}
 I(s; z_1) &= H(s) + H(z_1) - H(s, z_1) \\
 &= H(s) + H(z_1) - H(s | z_1) - H(z_1) \\
 &= H(s) - H(s | z_1)
 \end{aligned}$$

Then,

$$I(s; z_1, z_2) - I(s; z_1) = H(s | z_1) + H(z_2 | z_1) - H(s, z_2 | z_1)$$

which is the result.

Example 2.8 (other information) *This is a continuation of example 2.7 but with other information, z_2 , in the mix. Everything (including state prices) is the same as before except information z_2 is added with conditional probability of z_2 given transition state s and initial state z_1*

$\Pr(z_2 s, z_1)$	$z_2 = 0$	$z_2 = 1$
$s = 1, z_1 = 1$	0.4910536	0.5089464
$s = 1, z_1 = 2$	0.5035406	0.4964594
$s = 1, z_1 = 3$	0.5063793	0.4936207
$s = 2, z_1 = 1$	0.5050147	0.4949853
$s = 2, z_1 = 2$	0.4840381	0.5159619
$s = 2, z_1 = 3$	0.5046641	0.4953359
$s = 3, z_1 = 1$	0.5046469	0.4953531
$s = 3, z_1 = 2$	0.5077393	0.4922607
$s = 3, z_1 = 3$	0.4926152	0.5073848

*This implies the state-transition probabilities are equally likely to be*⁸

$$F_0 = \begin{bmatrix} 0.342567 & 0.366635 & 0.290798 \\ 0.378519 & 0.225076 & 0.396405 \\ 0.408194 & 0.169107 & 0.422698 \end{bmatrix}$$

or

$$F_1 = \begin{bmatrix} 0.35505 & 0.361479 & 0.283471 \\ 0.371002 & 0.23992 & 0.389078 \\ 0.400677 & 0.163952 & 0.435372 \end{bmatrix}$$

where

$$F = 0.5F_0 + 0.5F_1$$

$$\Pr(z_2 | z_1) = \Pr(z_2) = \begin{matrix} 0.5 & z_2 = 0 \\ 0.5 & z_2 = 1 \end{matrix}$$

⁸This is an atypical case in which other information z_2 is independent of initial state

information z_1 and the matrices F , F_0 , and F_1 are simultaneously diagonalizable (have the same eigenvectors) but different eigenvalues except they all share a maximum eigenvalue equal to one.

and the steady-state probability distribution for each of these state-transition distributions is

$$p_{ss}^T = [0.375877 \quad 0.257781 \quad 0.366342]$$

The joint distribution of states, s , and information, (z_1, z_2) , is

$$\begin{bmatrix} 0.0643816 & 0.0689048 & 0.0546522 \\ 0.0487875 & 0.0290101 & 0.0510928 \\ 0.0747694 & 0.0309755 & 0.0774261 \\ 0.0667275 & 0.0679358 & 0.0532752 \\ 0.0478186 & 0.0309234 & 0.0501484 \\ 0.0733924 & 0.0300312 & 0.0797475 \end{bmatrix}$$

the state distribution (sum of the columns) equals the steady-state distribution

$$\Pr(s) = p_{ss}$$

and the distribution for the information signals (sum of the rows) is

$$\Pr(z_1, z_2) = 0.5 \begin{bmatrix} p_1^{ss} \sum_{i=1}^3 F_{1i}^0 \\ p_2^{ss} \sum_{i=1}^3 F_{2i}^0 \\ p_3^{ss} \sum_{i=1}^3 F_{3i}^0 \\ p_1^{ss} \sum_{i=1}^3 F_{1i}^1 \\ p_2^{ss} \sum_{i=1}^3 F_{2i}^1 \\ p_3^{ss} \sum_{i=1}^3 F_{3i}^1 \end{bmatrix}$$

$$\Pr(z_1, z_2) = \begin{bmatrix} 0.187939 \\ 0.12889 \\ 0.183171 \\ 0.187939 \\ 0.12889 \\ 0.183171 \end{bmatrix}$$

Now, the expected gain from information, (z_1, z_2) , equals mutual information.

$$\begin{aligned} I(s; z_1, z_2) &= H(s) + H(z_1, z_2) - H(s, z_1, z_2) \\ &= 1.08513 + 1.77828 - 2.84256 \\ &= 0.0208427 \end{aligned}$$

or expected gain in growth rate equal to $\exp[0.0208427] = 1.02106$. This exceeds the expected gain from initial state information alone by the difference in mutual information.

$$\begin{aligned} I(s; z_1, z_2) - I(s; z_1) &= 0.0208427 - 0.0207388 \\ &= 0.000103864 \end{aligned}$$

or difference in expected growth rate equal to $\exp[0.000103864] = 1.0001$. For clarity, we report the expected returns from each information signal and relate them to $E[r | z_1, z_2]$ as well as

$$I(s; z_2 | z_1) = E[r | z_2, z_1] - E[r | z_1]$$

and

$$I(s; z_2 | z_1) = I(s; z_1, z_2) - I(s; z_1)$$

The long-run expected return conditional on (z_1, z_2) is as follows.

$$\begin{aligned} E[r_1 | z_2 = 0] &= [0.342567 \quad 0.366635 \quad 0.290798] \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.365559} & 0 \\ 0 & 0 & \frac{1}{0.302115} \end{bmatrix} \begin{bmatrix} 0.342567 \\ 0.366635 \\ 0.290798 \end{bmatrix} \right) \\ &= 0.000372432 \end{aligned}$$

$$\begin{aligned} E[r_2 | z_2 = 0] &= [0.378519 \quad 0.225076 \quad 0.396405] \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.338782} & 0 & 0 \\ 0 & \frac{1}{0.221511} & 0 \\ 0 & 0 & \frac{1}{0.392087} \end{bmatrix} \begin{bmatrix} 0.378519 \\ 0.225076 \\ 0.396405 \end{bmatrix} \right) \\ &= 0.0499156 \end{aligned}$$

$$\begin{aligned} E[r_3 | z_2 = 0] &= [0.408194 \quad 0.169107 \quad 0.422698] \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.348914} & 0 & 0 \\ 0 & \frac{1}{0.151415} & 0 \\ 0 & 0 & \frac{1}{0.408762} \end{bmatrix} \begin{bmatrix} 0.408194 \\ 0.169107 \\ 0.422698 \end{bmatrix} \right) \\ &= 0.0969123 \end{aligned}$$

$$\begin{aligned} E[r_1 | z_2 = 1] &= [0.35505 \quad 0.361479 \quad 0.283471] \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.365559} & 0 \\ 0 & 0 & \frac{1}{0.302115} \end{bmatrix} \begin{bmatrix} 0.35505 \\ 0.361479 \\ 0.283471 \end{bmatrix} \right) \\ &= 0.00137008 \end{aligned}$$

$$\begin{aligned} E[r_2 | z_2 = 1] &= [0.371002 \quad 0.23992 \quad 0.389078] \\ &\times \log \left(\begin{bmatrix} \frac{1}{0.338782} & 0 & 0 \\ 0 & \frac{1}{0.221511} & 0 \\ 0 & 0 & \frac{1}{0.392087} \end{bmatrix} \begin{bmatrix} 0.371002 \\ 0.23992 \\ 0.389078 \end{bmatrix} \right) \\ &= 0.0498609 \end{aligned}$$

$$\begin{aligned}
E[r_3 | z_2 = 1] &= [0.400677 \quad 0.163952 \quad 0.435372] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.348914} & 0 & 0 \\ 0 & \frac{1}{0.151415} & 0 \\ 0 & 0 & \frac{1}{0.408762} \end{bmatrix} \begin{bmatrix} 0.400677 \\ 0.163952 \\ 0.435372 \end{bmatrix} \right) \\
&= 0.0959249
\end{aligned}$$

Then,

$$\begin{aligned}
E[r | z_1, z_2] &= 0.5p_{ss}^T \left(\begin{bmatrix} E[r_1 | z_2 = 0] \\ E[r_2 | z_2 = 0] \\ E[r_3 | z_2 = 0] \end{bmatrix} + \begin{bmatrix} E[r_1 | z_2 = 1] \\ E[r_2 | z_2 = 1] \\ E[r_3 | z_2 = 1] \end{bmatrix} \right) \\
&= 0.0485099
\end{aligned}$$

or expected long-run growth rate equal to $\exp[0.0485099] = 1.04971$. Again, the expected gain from information (z_1, z_2) equals mutual information

$$\begin{aligned}
E[r | z_1, z_2] - E[r] &= I(s; z_1, z_2) \\
0.0485099 - 0.0276673 &= 0.0208427
\end{aligned}$$

or expected long-run gain in growth rate equals $\exp[0.0208427] = 1.02106$. Also, the long-run expected marginal gain from the additional information z_2 equals the difference in mutual information.

$$\begin{aligned}
(E[r | z_1, z_2] - E[r]) - (E[r | z_1] - E[r]) &= E[r | z_1, z_2] - E[r | z_1] \\
(0.0208427) - (0.0207388) &= 0.0485099 - 0.04840605 \\
&= 0.000103864 \\
I(s; z_2 | z_1) &= I(s; z_1, z_2) - I(s; z_1)
\end{aligned}$$

or difference in expected growth rate equal to $\exp[0.000103864] = 1.0001$.

The above example is a special case in which z_2 is independent of z_1 . The next example illustrates a more typical complementary scenario; z_2 is also conditionally more informative than the preceding case.

Example 2.9 (a more informative case) *Continue with the same setting as in example 2.8 except the conditional probability distributions of z_2 given transition state s and initial state z_1 are*

$\Pr(z_2 s, z_1)$	$z_2 = 0$	$z_2 = 1$
$s = 1, z_1 = 1$	0.7	0.3
$s = 1, z_1 = 2$	0.7	0.3
$s = 1, z_1 = 3$	0.3	0.7
$s = 2, z_1 = 1$	0.3	0.7
$s = 2, z_1 = 2$	0.3	0.7
$s = 2, z_1 = 3$	0.7	0.3
$s = 3, z_1 = 1$	0.7	0.3
$s = 3, z_1 = 2$	0.3	0.7
$s = 3, z_1 = 3$	0.7	0.3

This implies the state-transition probabilities are

$$F_0 = \begin{bmatrix} 0.4172736 & 0.43551468 & 0.1472118 \\ 0.2459615 & 0.15259206 & 0.6014465 \\ 0.4469691 & 0.07887546 & 0.4741554 \end{bmatrix}$$

or

$$F_1 = \begin{bmatrix} 0.2522396 & 0.2632663 & 0.4844941 \\ 0.4832028 & 0.2997742 & 0.2170230 \\ 0.3309515 & 0.3179672 & 0.3510812 \end{bmatrix}$$

and

$$\begin{aligned} F &= \begin{bmatrix} \Pr(z_2 = 0 | z_1 = 1) & 0 & 0 \\ 0 & \Pr(z_2 = 0 | z_1 = 2) & 0 \\ 0 & 0 & \Pr(z_2 = 0 | z_1 = 3) \end{bmatrix} F_0 \\ &+ \begin{bmatrix} \Pr(z_2 = 1 | z_1 = 1) & 0 & 0 \\ 0 & \Pr(z_2 = 1 | z_1 = 2) & 0 \\ 0 & 0 & \Pr(z_2 = 1 | z_1 = 3) \end{bmatrix} F_1 \\ &= \begin{bmatrix} 0.5851461 & 0 & 0 \\ 0 & 0.4570967 & 0 \\ 0 & 0 & 0.6333883 \end{bmatrix} F_0 \\ &+ \begin{bmatrix} 0.4148539 & 0 & 0 \\ 0 & 0.5429033 & 0 \\ 0 & 0 & 0.3666117 \end{bmatrix} F_1 \\ &= \begin{bmatrix} 0.348809 & 0.364057 & 0.287135 \\ 0.374761 & 0.232498 & 0.392742 \\ 0.404436 & 0.166529 & 0.429035 \end{bmatrix} \end{aligned}$$

The joint distribution of states, s , and information, (z_1, z_2) , is

$$\begin{bmatrix} 0.09177641 & 0.09578841 & 0.03237820 \\ 0.02898182 & 0.01798003 & 0.07086887 \\ 0.10371330 & 0.01830201 & 0.11002154 \\ 0.03933275 & 0.04105218 & 0.07554913 \\ 0.06762424 & 0.04195341 & 0.03037237 \\ 0.04444856 & 0.04270470 & 0.04715209 \end{bmatrix}$$

the state distribution (sum of the columns) equals the steady-state distribution

$$\Pr(s) = p_{ss}$$

and the distribution for the information signals (sum of the rows) is

$$\Pr(z_1, z_2) = \begin{bmatrix} 0.2199430 \\ 0.1178307 \\ 0.2320368 \\ 0.1559341 \\ 0.1399500 \\ 0.1343053 \end{bmatrix}$$

Now, the expected gain from information, (z_1, z_2) , equals mutual information.

$$\begin{aligned} I(s; z_1, z_2) &= H(s) + H(z_1, z_2) - H(s, z_1, z_2) \\ &= 1.08513 + 1.75865 - 2.76038 \\ &= 0.083399 \end{aligned}$$

or expected gain in growth rate equal to $\exp[0.083399] = 1.08698$. This exceeds the expected gain from initial state information alone by the difference in mutual information.

$$\begin{aligned} I(s; z_1, z_2) - I(s; z_1) &= 0.083399 - 0.020739 \\ &= 0.062660 \end{aligned}$$

or difference in expected growth rate equal to $\exp[0.062660] = 1.06467$. Again for completeness, we report the expected returns from each information signal and $E[r | z_1, z_2]$ then relate them to

$$I(s; z_2 | z_1) = E[r | z_2, z_1] - E[r | z_1]$$

and

$$I(s; z_2 | z_1) = I(s; z_1, z_2) - I(s; z_1)$$

The long-run expected return conditional on (z_1, z_2) is as follows.

$$\begin{aligned} E[r_1 | z_2 = 0] &= [0.4172736 \quad 0.43551468 \quad 0.1472118] \\ &\times \log \left(\left[\begin{array}{ccc} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.365559} & 0 \\ 0 & 0 & \frac{1}{0.302115} \end{array} \right] \left[\begin{array}{c} 0.4172736 \\ 0.43551468 \\ 0.1472118 \end{array} \right] \right) \\ &= 0.06540519 \end{aligned}$$

$$\begin{aligned} E[r_2 | z_2 = 0] &= [0.2459615 \quad 0.15259206 \quad 0.6014465] \\ &\times \log \left(\left[\begin{array}{ccc} \frac{1}{0.338782} & 0 & 0 \\ 0 & \frac{1}{0.221511} & 0 \\ 0 & 0 & \frac{1}{0.392087} \end{array} \right] \left[\begin{array}{c} 0.2459615 \\ 0.15259206 \\ 0.6014465 \end{array} \right] \right) \\ &= 0.1217062 \end{aligned}$$

$$\begin{aligned} E[r_3 | z_2 = 0] &= [0.4469691 \quad 0.0788755 \quad 0.4741554] \\ &\times \log \left(\left[\begin{array}{ccc} \frac{1}{0.348914} & 0 & 0 \\ 0 & \frac{1}{0.151415} & 0 \\ 0 & 0 & \frac{1}{0.408762} \end{array} \right] \left[\begin{array}{c} 0.4469691 \\ 0.0788755 \\ 0.4741554 \end{array} \right] \right) \\ &= 0.1296252 \end{aligned}$$

$$\begin{aligned}
E[r_1 | z_2 = 1] &= [0.2522396 \quad 0.2632663 \quad 0.4844941] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.365559} & 0 \\ 0 & 0 & \frac{1}{0.302115} \end{bmatrix} \begin{bmatrix} 0.2522396 \\ 0.2632663 \\ 0.4844941 \end{bmatrix} \right) \\
&= 0.07285352
\end{aligned}$$

$$\begin{aligned}
E[r_2 | z_2 = 1] &= [0.4832028 \quad 0.2997742 \quad 0.2170230] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.338782} & 0 & 0 \\ 0 & \frac{1}{0.221511} & 0 \\ 0 & 0 & \frac{1}{0.392087} \end{bmatrix} \begin{bmatrix} 0.4832028 \\ 0.2997742 \\ 0.2170230 \end{bmatrix} \right) \\
&= 0.1339083
\end{aligned}$$

$$\begin{aligned}
E[r_3 | z_2 = 1] &= [0.3309515 \quad 0.3179672 \quad 0.3510812] \\
&\times \log \left(\begin{bmatrix} \frac{1}{0.348914} & 0 & 0 \\ 0 & \frac{1}{0.151415} & 0 \\ 0 & 0 & \frac{1}{0.408762} \end{bmatrix} \begin{bmatrix} 0.3309515 \\ 0.3179672 \\ 0.3510812 \end{bmatrix} \right) \\
&= 0.1650103
\end{aligned}$$

Then,

$$\begin{aligned}
E[r | z_1, z_2] &= \Pr(z_1, z_2)^T \begin{bmatrix} E[r_1 | z_2 = 0] \\ E[r_2 | z_2 = 0] \\ E[r_3 | z_2 = 0] \\ E[r_1 | z_2 = 1] \\ E[r_2 | z_2 = 1] \\ E[r_3 | z_2 = 1] \end{bmatrix} \\
&= 0.1110666
\end{aligned}$$

or expected long-run growth rate equal to $\exp[0.1110666] = 1.117469$. Again, the expected gain from information (z_1, z_2) equals mutual information

$$\begin{aligned}
E[r | z_1, z_2] - E[r] &= I(s; z_1, z_2) \\
0.1110666 - 0.0276673 &= 0.0833993
\end{aligned}$$

or expected long-run gain in growth rate equals $\exp[0.0833993] = 1.086976$. Also, the long-run expected marginal gain from the additional information z_2 equals the difference in mutual information.

$$\begin{aligned}
(E[r | z_1, z_2] - E[r]) - (E[r | z_1] - E[r]) &= E[r | z_1, z_2] - E[r | z_1] \\
(0.0833993) - (0.0207388) &= 0.1110666 - 0.04840605 \\
&= 0.062660 \\
I(s; z_2 | z_1) &= I(s; z_1, z_2) - I(s; z_1)
\end{aligned}$$

or difference in expected growth rate equal to $\exp[0.062660] = 1.06467$.

2.6 Joint distribution from full set of conditional distributions

In the foregoing analyses we've derived the joint distribution associated with initial and post-transition states by utilizing the steady-state distribution in conjunction with the state-transition (state at $t + 1$) distribution conditional on the initial state (state at t), F . In this section, we offer another but equivalent approach illustrating the power of the recovery theorem.

Since Markov processes are reversible,⁹ we can utilize the recovery theorem frame to derive the reverse state-transition (state at $t - 1$ conditional on state at t) distribution. Label this conditional distribution G . Together F and G represent the full set of conditional distributions from which we can derive the joint distribution (Besag [1974]).

The reverse state-transition distribution utilizes P^T in place of P

$$G = \frac{1}{\delta} D_G P^T D_G^{-1}$$

where eigensystem decomposition of P^T involves

$$\begin{aligned} P^T \zeta_G &= \delta \zeta_G \\ P^T \zeta_G &= P^T D_G^{-1} \iota \end{aligned}$$

D_G^{-1} is a diagonal matrix composed of the elements of the eigenvector ζ_G and G_{ij} refers to transitioning from state i at t to state j at $t - 1$.¹⁰

The joint distribution can be derived from the full set of conditional distributions. Let s_0 equal the state at t and s_1 the state at $t + 1$ for one set of draws and t_0 and t_1 the analogous states for another set of draws. The joint distribution is

$$\Pr(s_0, s_1) = \Pr(s_1 | s_0) \Pr(s_0)$$

The first term on the right hand side is a row from F , but the second term, the marginal distribution $p(s_0)$, is unknown. However,

$$\Pr(s_0) = \frac{\Pr(s_0, t_1)}{\Pr(t_1 | s_0)}$$

therefore,

$$\Pr(s_0, s_1) = \frac{\Pr(s_1 | s_0) \Pr(s_0, t_1)}{\Pr(t_1 | s_0)}$$

⁹Fundamentally, there is no substantive distinction between forecasting and "backcasting".

¹⁰Diagonalization of P can be written

$$P = S \Lambda S^{-1}$$

where S is a matrix of eigenvectors and Λ is a diagonal matrix of eigenvalues. Then,

$$P^T = (S^{-1})^T \Lambda S^T$$

This leaves $\Pr(s_0, t_1)$. Repeating a similar strategy leads to¹¹

$$\Pr(s_0, t_1) = \frac{\Pr(s_0 | t_1) \Pr(t_0, t_1)}{\Pr(t_0 | t_1)}$$

Again, substitute

$$\Pr(s_0, s_1) = \frac{\Pr(s_1 | s_0) \Pr(s_0 | t_1) \Pr(t_0, t_1)}{\Pr(t_1 | s_0) \Pr(t_0 | t_1)}$$

and we have the ratio of unknown joint likelihoods expressed in terms of the ratios of known conditional likelihoods.

$$\frac{\Pr(s_0, s_1)}{\Pr(t_0, t_1)} = \frac{\Pr(s_1 | s_0) \Pr(s_0 | t_1)}{\Pr(t_1 | s_0) \Pr(t_0 | t_1)}$$

In the recovery theorem frame, this can be expressed

$$\frac{\Pr(s_0 = i, s_1 = j)}{\Pr(t_0 = k, t_1 = \ell)} = \frac{F_{ij} G_{\ell i}}{F_{i\ell} G_{\ell k}}$$

Since we can derive the relative joint likelihoods and the sum of the joint likelihoods equals one, we have a simple linear system from which the joint probability assignment can be deduced. Next, we illustrate these ideas by returning to example 2.7.

Example 2.10 (joint dynamic distribution) *This is a continuation of example 2.7. Recall,*

$$\begin{aligned} F &= \frac{1}{0.952747} \begin{bmatrix} 1.64931 & 0 & 0 \\ 0 & 1.73825 & 0 \\ 0 & 0 & 1.82142 \end{bmatrix} \\ &\times \begin{bmatrix} 0.332326 & 0.365559 & 0.302115 \\ 0.338782 & 0.221511 & 0.392087 \\ 0.348914 & 0.151415 & 0.408762 \end{bmatrix} \begin{bmatrix} 0.606314 & 0 & 0 \\ 0 & 0.57529 & 0 \\ 0 & 0 & 0.549021 \end{bmatrix} \\ &= \begin{bmatrix} 0.348809 & 0.364057 & 0.287135 \\ 0.374761 & 0.232498 & 0.392742 \\ 0.404436 & 0.166529 & 0.429035 \end{bmatrix} \end{aligned}$$

Steady-state probabilities are

$$p_{ss}^T F = p_{ss}^T$$

¹¹To verify this claim, notice

$$\begin{aligned} \Pr(s_0, t_1) &= \frac{\Pr(s_0 | t_1) \Pr(t_0, t_1)}{\Pr(t_0 | t_1)} \times \frac{\Pr(t_1)}{\Pr(t_1)} \\ &= \frac{\Pr(s_0, t_1) \Pr(t_0, t_1)}{\Pr(t_0, t_1)} \\ &= \Pr(s_0, t_1) \end{aligned}$$

$$p_{ss}^T = [0.375877 \quad 0.257781 \quad 0.366342]$$

Then,

$$\begin{aligned} G &= \frac{1}{0.952747} \begin{bmatrix} 1.61307 & 0 & 0 \\ 0 & 2.2317 & 0 \\ 0 & 0 & 1.49866 \end{bmatrix} \\ &\times \begin{bmatrix} 0.332326 & 0.365559 & 0.302115 \\ 0.338782 & 0.221511 & 0.392087 \\ 0.348914 & 0.151415 & 0.408762 \end{bmatrix}^T \begin{bmatrix} 0.619938 & 0 & 0 \\ 0 & 0.448088 & 0 \\ 0 & 0 & 0.667264 \end{bmatrix} \\ &= \begin{bmatrix} 0.348809 & 0.257015 & 0.394176 \\ 0.530841 & 0.232498 & 0.236661 \\ 0.294608 & 0.276357 & 0.429035 \end{bmatrix} \end{aligned}$$

The (same) steady-state distribution can be derived as well from G .

$$p_{ss}^T G = p_{ss}^T$$

$$p_{ss}^T = [0.375877 \quad 0.257781 \quad 0.366342]$$

With the full set of conditional distributions, we can now derive the joint probability distribution of the state at t and the state at $t+1$. We employ $\Pr(s_0 = 1, s_1 = 1)$ as the reference joint likelihood and develop all likelihood ratios with it as the numerator.

$$\begin{aligned} \frac{\Pr(s_0 = 1, s_1 = 1)}{\Pr(t_0 = 1, t_1 = 2)} &= \frac{F_{11}G_{21}}{F_{12}G_{21}} \\ &= \frac{(0.348809)(0.530841)}{(0.364057)(0.530841)} \\ &= 0.958116 \end{aligned}$$

$$\begin{aligned} \frac{\Pr(s_0 = 1, s_1 = 1)}{\Pr(t_0 = 1, t_1 = 3)} &= \frac{F_{11}G_{31}}{F_{13}G_{31}} \\ &= \frac{(0.348809)(0.294608)}{(0.287135)(0.294608)} \\ &= 1.21479 \end{aligned}$$

$$\begin{aligned} \frac{\Pr(s_0 = 1, s_1 = 1)}{\Pr(t_0 = 2, t_1 = 1)} &= \frac{F_{11}G_{11}}{F_{11}G_{12}} \\ &= \frac{(0.348809)(0.348809)}{(0.348809)(0.257015)} \\ &= 1.35715 \end{aligned}$$

$$\begin{aligned} \frac{\Pr(s_0 = 1, s_1 = 1)}{\Pr(t_0 = 2, t_1 = 2)} &= \frac{F_{11}G_{21}}{F_{12}G_{22}} \\ &= \frac{(0.348809)(0.530841)}{(0.364057)(0.232498)} \\ &= 2.18758 \end{aligned}$$

$$\begin{aligned}
\frac{\Pr(s_0 = 1, s_1 = 1)}{\Pr(t_0 = 2, t_1 = 3)} &= \frac{F_{11}G_{31}}{F_{13}G_{32}} \\
&= \frac{(0.348809)(0.294608)}{(0.287135)(0.276357)} \\
&= 1.29502
\end{aligned}$$

$$\begin{aligned}
\frac{\Pr(s_0 = 1, s_1 = 1)}{\Pr(t_0 = 3, t_1 = 1)} &= \frac{F_{11}G_{11}}{F_{11}G_{13}} \\
&= \frac{(0.348809)(0.348809)}{(0.348809)(0.394176)} \\
&= 0.884905
\end{aligned}$$

$$\begin{aligned}
\frac{\Pr(s_0 = 1, s_1 = 1)}{\Pr(t_0 = 3, t_1 = 2)} &= \frac{F_{11}G_{21}}{F_{12}G_{23}} \\
&= \frac{(0.348809)(0.530841)}{(0.364057)(0.236661)} \\
&= 2.14909
\end{aligned}$$

$$\begin{aligned}
\frac{\Pr(s_0 = 1, s_1 = 1)}{\Pr(t_0 = 3, t_1 = 3)} &= \frac{F_{11}G_{31}}{F_{13}G_{33}} \\
&= \frac{(0.348809)(0.294608)}{(0.287135)(0.429035)} \\
&= 0.834168
\end{aligned}$$

Now, we solve

$$Ap = b$$

where¹²

$$A = \begin{bmatrix} 1 & -0.958 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1.214 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1.357 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -2.187 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1.295 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -0.884 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & -2.149 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.834 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

¹²We have dropped some digits in A (relative to that reported earlier) to better fit the space.

$$p = \begin{bmatrix} \Pr(s_0 = 1, s_1 = 1) \\ \Pr(s_0 = 1, s_1 = 2) \\ \Pr(s_0 = 1, s_1 = 3) \\ \Pr(s_0 = 2, s_1 = 1) \\ \Pr(s_0 = 2, s_1 = 2) \\ \Pr(s_0 = 2, s_1 = 3) \\ \Pr(s_0 = 3, s_1 = 1) \\ \Pr(s_0 = 3, s_1 = 2) \\ \Pr(s_0 = 3, s_1 = 3) \end{bmatrix}$$

and

$$b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The solution is

$$\begin{bmatrix} \Pr(s_0 = 1, s_1 = 1) \\ \Pr(s_0 = 1, s_1 = 2) \\ \Pr(s_0 = 1, s_1 = 3) \\ \Pr(s_0 = 2, s_1 = 1) \\ \Pr(s_0 = 2, s_1 = 2) \\ \Pr(s_0 = 2, s_1 = 3) \\ \Pr(s_0 = 3, s_1 = 1) \\ \Pr(s_0 = 3, s_1 = 2) \\ \Pr(s_0 = 3, s_1 = 3) \end{bmatrix} = \begin{bmatrix} 0.131109 \\ 0.136841 \\ 0.107927 \\ 0.0966061 \\ 0.0599334 \\ 0.101241 \\ 0.148162 \\ 0.0610067 \\ 0.157174 \end{bmatrix}$$

This matches the joint distribution reported earlier which was derived via

$$[F_1 p_1^{ss}, F_2 p_2^{ss}, F_3 p_3^{ss}]$$

The Kelly-Ross theorem applies to reverse probability assignment as well as forward assignment, therefore the long-run expected growth rate equals $\frac{1}{\delta} = 1.0496$. To support this claim, we enumerate state-by-state expected returns and growth

rates. For initial state one, the long-run expected return is

$$\begin{aligned}
E[r_1^G | z] &= G_1 \log(\Omega_1^G G_1^T) \\
&= [0.348809 \quad 0.257015 \quad 0.394176] \\
&\quad \times \log \left(\begin{bmatrix} \frac{1}{0.332326} & 0 & 0 \\ 0 & \frac{1}{0.338782} & 0 \\ 0 & 0 & \frac{1}{0.348914} \end{bmatrix} \begin{bmatrix} 0.348809 \\ 0.257015 \\ 0.394176 \end{bmatrix} \right) \\
&= [0.348809 \quad 0.257015 \quad 0.394176] \log \begin{bmatrix} 1.0496 \\ 0.758644 \\ 1.12972 \end{bmatrix} \\
&= -0.0060299
\end{aligned}$$

or the long-run expected growth rate equals $\exp[-0.0060299] = 0.993988$. For initial state two, the long-run expected return is

$$\begin{aligned}
E[r_2^G | z] &= G_2 \log(\Omega_2^G G_2^T) \\
&= [0.530841 \quad 0.232498 \quad 0.236661] \\
&\quad \times \log \left(\begin{bmatrix} \frac{1}{0.365559} & 0 & 0 \\ 0 & \frac{1}{0.221511} & 0 \\ 0 & 0 & \frac{1}{0.151415} \end{bmatrix} \begin{bmatrix} 0.530841 \\ 0.232498 \\ 0.236661 \end{bmatrix} \right) \\
&= [0.530841 \quad 0.232498 \quad 0.236661] \log \begin{bmatrix} 1.45214 \\ 1.0496 \\ 1.56299 \end{bmatrix} \\
&= 0.31497
\end{aligned}$$

or the long-run expected growth rate equals $\exp[0.31497] = 1.37022$. For initial state three, the long-run expected return is

$$\begin{aligned}
E[r_3^G | z] &= G_3 \log(\Omega_3^G G_3^T) \\
&= [0.294608 \quad 0.276357 \quad 0.429035] \\
&\quad \times \log \left(\begin{bmatrix} \frac{1}{0.365559} & 0 & 0 \\ 0 & \frac{1}{0.221511} & 0 \\ 0 & 0 & \frac{1}{0.151415} \end{bmatrix} \begin{bmatrix} 0.294608 \\ 0.276357 \\ 0.429035 \end{bmatrix} \right) \\
&= [0.294608 \quad 0.276357 \quad 0.429035] \log \begin{bmatrix} 0.975152 \\ 0.704836 \\ 1.0496 \end{bmatrix} \\
&= -0.083312
\end{aligned}$$

or the long-run expected growth rate equals $\exp[-0.083312] = 0.920064$. Hence,

the expected long-run logarithmic return is

$$\begin{aligned}
 E[r^G | z] &= p_{ss}^T \begin{bmatrix} E[r_1^G | z] \\ E[r_2^G | z] \\ E[r_3^G | z] \end{bmatrix} \\
 &= \begin{bmatrix} 0.375877 & 0.257781 & 0.366342 \end{bmatrix} \begin{bmatrix} -0.0060299 \\ 0.31497 \\ -0.083312 \end{bmatrix} \\
 &= 0.048406
 \end{aligned}$$

or the expected long-run growth rate equals $\exp[0.048406] = 1.0496 = \frac{1}{\delta}$.

3 Discussion

We have discussed and illustrated information synergy. The marriage of recovery theorem probability assignment and a Kelly investment strategy highlights the importance of covering all bases (spanning the state space) to guard against bankruptcy ruin in the long-run. While maximization of long-run wealth or the expected growth rate easily translates into a financial wealth perspective, the idea applies more broadly to anything valued — that is, utility. Organizational synergy, broadly construed, is enhanced via careful attention to what is known in assigning likelihoods and diversifying so that states are spanned. In other words, this implies organizational complementarities are vital to the long-run.

Unfortunately, short-run focused performance rewards may discourage building complementarities.¹³ The omnipresent dwelling on short-run performance appears ironic as accounting is often the source of performance measurement and though accounting is effective as a long-run score card it suffers many deficiencies as a short-run performance measure. With its ability to map likelihoods into returns (seemingly entirely different quantities), the mutual information theorem affords accounting an opportunity to convey long-run performance information without revealing proprietary details. Accounting supplies events data that can be utilized to update state beliefs and, if allowed time, sound long-run performance measures as short-run deviations are settled up over the long term.

Future work (parts 2 and 3) involves exploration of the implications of finer short-term focussed other information, coarser long-term focussed accounting information, and their combined information. Fine information may be self-reinforcing but lead economic agents astray from their long-term wealth goals. Coarse information may be relatively slow to adapt but eventually provide a compass toward long-term wealth creation.

¹³In game theory terms, contrast one-shot games or known end-point games with repeated games or unknown end-point games to which the folk theorems apply.