

DESEMPEÑO FINANCIERO DE FONDOS DE PENSIONES*

Salvador Zurita y Carlos Jara

A pesar de la creciente importancia de la industria previsional en la economía chilena, las mediciones de desempeño financiero suelen ser bastante simples, consistiendo sólo en retornos promedios en períodos de uno y tres años. Los objetivos del artículo son: (1) revisar la literatura de medición de desempeño financiero de portafolios de inversión, y proponer una medida de desempeño ajustado por riesgo aplicable a los fondos de pensiones chilenos, (2) comparar el desempeño de los fondos de pensiones según su desempeño ajustado por riesgo en el período enero de 1987 a septiembre de 1998, y (3) estudiar la persistencia del desempeño de las AFPs en el tiempo, por sus implicancias sobre la decisión de cambiarse de administradora.

Las conclusiones que aquí se presentan son las siguientes:

a) el índice de Sharpe tiene ventajas como medida de desempeño ajustada por riesgo para los fondos de pensiones porque (i) supone

SALVADOR ZURITA. Ingeniero Comercial, Universidad de Chile. Ph. D. en Finanzas, Universidad de California, Los Ángeles. Profesor e Investigador, Universidad de Chile.

CARLOS JARA. Licenciado en Matemática, Universidad de Santiago. Magister en Administración (mención Finanzas), Universidad de Chile. Profesor y Consultor de Empresas.

* Los autores agradecen los comentarios de Patricio Arrau, Harald Beyer y participantes de seminarios realizados en el CEP y en el Encuentro Nacional de Facultades de Administración de 1998. Una versión preliminar de este trabajo se publicó en la serie *Documentos de Trabajo*, N° 276, enero 1998.

que los fondos de la cuenta administrada por las AFPs representa gran parte de la riqueza de los cotizantes al jubilar, lo que parece apropiado para la mayoría de los pensionados a la fecha de jubilación, (ii) no depende de un modelo específico de valoración de activos, y (iii) puesto que en particular no se basa en el CAPM, no requiere identificar el portafolio de mercado.

b) utilizando períodos de 3 años para medir, los rankings de desempeño según rentabilidad promedio (medida actual) y Sharpe están correlacionados positivamente, pero presentan diferencias importantes en algunos períodos. En efecto, el coeficiente de correlación (Spearman) entre rankings basados en rentabilidad promedio (criterio actual) y ajustados por riesgo (Sharpe) fluctúa entre 0,45 y 0,98. Esta diferencia se atenúa si se consideran períodos de cinco años, en cuyo caso las correlaciones no fueron inferiores a 0,85 en ningún período.

c) se aprecia alguna evidencia de la racionalidad de cambiarse de AFP en base al desempeño histórico del fondo, pero ella no es concluyente. En efecto, al comparar los desempeños ajustados por riesgo entre períodos no superpuestos, se observa que en general las correlaciones seriales son positivas (aunque estadísticamente no distintas de cero sólo en un 20% de los casos).

Entre las empresas que administran fondos de terceros en Chile, ninguna es más importante que las administradoras de fondos de pensiones (AFP), las que a abril de 1998 administraban unos 30 billones de dólares, aproximadamente equivalente a un 39,4% del PIB estimado de 1998. A pesar de ello, los resultados de desempeño financiero que la Superintendencia hace públicos periódicamente son notoriamente simples, consistiendo sólo en retornos promedios de la cartera en períodos de uno y tres años. Tales mediciones ignoran otras características importantes de los portafolios, entre las que destaca el riesgo. Ello ha contribuido a que las AFPs tiendan a mantener carteras parecidas entre sí, porque si alguna escoge invertir en un portafolio de riesgo diferente, se expone a quedar por debajo del promedio, y eventualmente a cubrir la diferencia para alcanzar la rentabilidad mínima.

El propósito de este artículo es (i) revisar la literatura de medición de desempeño financiero de portafolios de inversión y su aplicabilidad al sistema de pensiones chileno, (ii) ordenar las administradoras por desempeño financiero ajustado por riesgo, en el período enero de 1987 a septiembre de 1998, comparando este ranking con uno basado en el promedio, y (iii)

discutir la racionalidad de cambiarse de AFP basado en desempeño histórico de las administradoras, en lo que se refiere a persistencia de los retornos en el tiempo. La sección que sigue contiene la revisión y discusión de la literatura, la sección II describe los datos y resultados empíricos, y la sección III resume y concluye.

I. MEDIDAS DE DESEMPEÑO DE CARTERAS DE INVERSIÓN

Medir el desempeño de un portafolio es comparar su retorno con los que el cliente podría haber obtenido si hubiese elegido uno o más portafolios alternativos, llamados portafolios *benchmark*. Idealmente un portafolio *benchmark* debe cumplir el doble requisito de ser una comparación relevante (uno que podría haber sido elegido por el inversionista), y alcanzable (debe incorporar restricciones similares a las del portafolio bajo evaluación). En general, hay dos formas de definir portafolios *benchmark* que satisfagan los requisitos anteriores: se pueden restringir las comparaciones a portafolios de riesgo similar (ésta ha sido la metodología en la industria en Chile), o bien se puede calcular una medida única ajustada por riesgo, permitiendo así comparar las carteras entre sí o con un *benchmark*, para diferentes niveles de riesgo. A continuación, revisamos las medidas ajustadas por riesgo propuestas en la literatura.

A. Medidas basadas en el modelo de media-varianza

En el modelo de media-varianza o modelo de dos parámetros surgen dos conceptos de riesgo: riesgo total (medido por la desviación estándar o por la varianza del retorno), y riesgo de mercado, sistemático, de covarianza o no diversificable (medido por el coeficiente beta). El primer concepto es apropiado si el activo o portafolio bajo evaluación representa la única riqueza del inversionista, pero si éste mantiene muchos activos el riesgo de covarianza es el relevante. Como en el Modelo de Valoración de Activos de Capital (CAPM) los inversionistas mantienen un único portafolio de activos riesgosos (el de mercado), el riesgo de covarianza se mide con respecto a él y se llama beta. En la literatura se han propuesto diferentes medidas de desempeño, basadas en estos conceptos de riesgo.

1. Coeficiente alfa de Jensen (1968)

La línea característica de un portafolio se define por¹:

$$r_{pt} - r_{lt} = \alpha_p + \beta_p (r_{mt} - r_{lt}) + \varepsilon_{pt}$$

Donde r_{pt} = retorno del portafolio p en el período t, r_{lt} = tasa de interés libre de riesgo en el período t, r_{mt} = retorno del portafolio de mercado en el período t, α_p y β_p = coeficientes, y ε_{pt} = error aleatorio.

La ecuación indica que el exceso de retorno del portafolio p (por sobre la tasa de interés libre de riesgo) en el período t se puede descomponer en tres partes: el alfa del portafolio, el premio por el riesgo (retorno en exceso del portafolio de mercado multiplicado por el beta del portafolio), y un término de error aleatorio representativo de riesgo diversificable.

Si los coeficientes alfa y beta se suponen constantes en el intervalo de evaluación, la ecuación anterior es lineal, y puede ser estimada por mínimos cuadrados ordinarios. Según el modelo de valoración de activos de capital (CAPM), en equilibrio todos los activos o portafolios deberían estar sobre la línea de mercado de activos, lo que implica alfa igual a cero. Luego, si el coeficiente alfa estimado resulta mayor que cero indica desempeño superior, y si menor que cero, inferior. Como alfa sólo puede ser estimada con error, se realiza un test t para determinar si la diferencia entre el alfa estimado y cero (hipótesis nula) es significativa o no.

2. Coeficiente de Treynor (1965)

Esta medida también se basa en la línea de mercado de activos (LMA) para definir un *benchmark*, y consiste en dividir el retorno en exceso promedio por el riesgo de mercado del portafolio:

$$Tp = \frac{ar_p - ar_l}{\beta_p}$$

donde ar_p = retorno promedio del portafolio p en el período de evaluación, ar_l = tasa de interés libre de riesgo promedio durante el período de evaluación, y β_p = riesgo sistemático del portafolio p.

¹ En valor esperado, corresponde a la línea de mercado de activos (LMA).

Así, corresponde a la pendiente de la línea que se origina en la tasa libre de riesgo promedio y pasa por el punto (β_p, ar_p) . El *benchmark* para comparar es la pendiente de la línea de mercado de activos expost, es decir, la línea que pasa por los puntos $(0, ar_l)$ y $(1, ar_m - ar_l)$. Es decir, el punto de comparación es el premio por riesgo expost $(ar_m - ar_l)$. Si T_p es mayor que este valor, el portafolio ha tenido un mejor desempeño que el mercado, y si menor, inferior.

Es fácil verificar que las dos medidas anteriores a_p y T_p basadas en la LMA expost, siempre dan la misma respuesta del desempeño del portafolio relativo al portafolio de mercado: cualquier portafolio con alfa expost positivo está por arriba de la LMA expost (indicación de desempeño superior según Jensen), y luego debe tener una pendiente mayor que la de la LMA expost (también una indicación de desempeño superior según Treynor), y similarmente para alfa negativo.

Sin embargo, es posible que estas dos medidas ordenen (ranking) los portafolios en forma diferente, simplemente porque los cálculos son distintos. Específicamente, es posible que dos portafolios A y B con riesgos sistemáticos β_A, β_B tales que $\beta_A < \beta_B$, presenten un $\alpha_A < \alpha_B$, indicando que B tuvo un desempeño superior según Jensen, pero simultáneamente, indicando que A tuvo desempeño superior según Treynor.

En nuestro caso, si los betas de los fondos son similares (como podría esperarse), el ranking de fondos según desempeño según Jensen y Treynor debiera tender a coincidir.

Finalmente, Dobson y Korkie (1981) proveen distribuciones asintóticas del estimador de Treynor, pero lamentablemente no se comporta bien en muestras pequeñas y además carecen de poder.

3. Coeficiente de Sharpe (1966)

A diferencia de las medidas de Jensen y Treynor que se basan en la LMA expost, el índice de Sharpe se basa en la línea de mercado de capitales (LMC) expost en el plano retorno medio – desviación estándar:

$$ar_p^e = ar_l + \frac{ar_m - ar_l}{\sigma_m} \sigma_p$$

Donde el superíndice e denota equilibrio, y σ_p, σ_m denotan las desviaciones estándar de los retornos del portafolio p y del portafolio de mercado, respectivamente.

Después de determinar la posición de la LMC expost, el retorno promedio y la desviación estándar del portafolio bajo evaluación se comparan contra ella. El índice de Sharpe se define como la pendiente de la línea que se origina en la tasa libre de riesgo promedio, y que pasa por el punto correspondiente a la desviación estándar y el retorno promedio del portafolio:

$$S_p = \frac{ar_p - ar_l}{\sigma_p}$$

Si este coeficiente resulta mayor que la pendiente de la LMC expost, ello indica desempeño superior, y si resulta menor indica desempeño inferior.

Al comparar los índices de Treynor y Sharpe se concluye que en ciertas situaciones pueden generar evaluaciones discrepantes con respecto a si el desempeño fue superior o inferior al mercado. Específicamente, es posible que el índice de Treynor indique desempeño superior al mercado, y el de Sharpe desempeño inferior, porque el portafolio puede tener una gran cantidad de riesgo específico, lo que no afecta el índice de Treynor (que sólo incluye el riesgo de mercado en el denominador) pero sí el de Sharpe (puesto que esta medida se basa en riesgo total). Sin embargo, en el caso de caso de portafolios bien diversificados por definición no hay riesgo residual, y por ello ambos criterios debieran dar la misma respuesta.

Finalmente, Dobson y Korkie (1981) desarrollaron distribuciones asintóticas para la medida de Sharpe, consistentes en un estadístico z para comparaciones simples, y en un estadístico chi-cuadrado para comparaciones múltiples, obteniendo un buen comportamiento en muestras pequeñas aunque el poder para detectar diferencias típicas con datos mensuales es pequeño.

4. Timing del mercado

Las tres medidas anteriores de desempeño permitirían detectar desempeño superior atribuible a selectividad, esto es, habilidad para detectar activos incorrectamente valorados por el mercado. Sin embargo, un administrador también puede ser hábil en anticipar los movimientos futuros del mercado (*market timing*), y ajustar su portafolio para tener un beta alto durante alzas de mercado y uno bajo durante bajas de mercado. El problema es que si el administrador es un 'timer' activo, las medidas anteriores de performance estarán sesgadas, puesto que el riesgo por el que se ajustan no es estable en el período de evaluación (Fama (1972) y Jensen (1972)).

Para corregir este problema se han propuesto diversos métodos. Treynor y Mazuy (1966) y Admati, Bhattacharya, Pfleiderer y Ross (1986) proponen el método de regresión cuadrática, que consiste en ajustar una curva característica en vez de una línea recta:

$$r_{pt} - r_{lt} = a + b(r_{mt} - r_{lt}) + c(r_{mt} - r_{lt})^2 + \varepsilon_{pt}$$

Donde ε_{pt} es un término de error aleatorio. Si el valor estimado de c es positivo (curva convexa), la curva tiene menos pendiente al moverse hacia la izquierda. Ello indica que el administrador tuvo éxito en anticipar el mercado. Si $c = 0$, a y b corresponden al alfa y beta del portafolio, respectivamente. Un coeficiente a positivo es indicativo de desempeño superior atribuible a selectividad (habilidad para detectar activos incorrectamente preciado). Este método supone que el administrador altera muchas veces el beta de su portafolio en el tiempo (la pendiente de la curva, creciente, representa el beta del portafolio).

Hendrikson y Merton (1981) sugieren un procedimiento alternativo, consistente en correr dos regresiones lineales, una para períodos cuando los activos riesgosos tuvieron mejor retorno que los libres de riesgo (mercados en alza, $r_m > r_f$), y otra cuando no tuvieron tan buen desempeño (mercados de baja, $r_m < r_f$). Un "timer" exitoso elegirá un alto beta durante mercados en alza y uno bajo en mercados en baja. Gráficamente, la pendiente de la línea característica para retornos en exceso positivos expost es mayor que la pendiente con retornos en exceso negativos. Este método requiere estimar

$$r_{pt} - r_{lt} = a + b(r_{mt} - r_{lt}) + c[D_t(r_{mt} - r_{lt})] + \varepsilon_t$$

Donde ε_t es un error aleatorio, y D_t es una variable muda que vale cero en cualquier período t en que $r_{mt} > r_{lt}$, y vale -1 en cualquier período t en que $r_{mt} < r_{lt}$.

Bhattacharya y Pfleiderer (1983) obtienen una relación similar a la de Treynor y Mazuy en términos de variables observadas, pero más refinada, pues analizan el término de error para identificar las habilidades de predicción del administrador. En su modelo suponen que los administradores asignan la probabilidad de que el mercado esté por sobre o por debajo de la tasa de interés libre de riesgo, dependiendo de la magnitud en que el retorno de mercado sea distinto a la tasa de interés. En este aspecto, el modelo es más sofisticado que el de Hendrikson y Merton (1981), quienes

suponen que los administradores sólo reciben señales referente a si el mercado estará por sobre o por debajo de la tasa de interés libre de riesgo, y no respecto a la magnitud de la diferencia.

Kon y Jen (1978) proponen utilizar el método estadístico de “regresión cambiante” para estimar el riesgo sistemático que varía en el tiempo y el desempeño de administradoras de fondos. El método consiste en estimar la medida de Jensen pero considerando el coeficiente de riesgo sistemático beta del portfolio como una variable aleatoria dentro del modelo, el cual puede tomar uno de dos valores, llamados regímenes (cada régimen corresponde a una línea de Jensen diferente, con distintos parámetros). Se hace el supuesto que el administrador elige el régimen 1 con probabilidad λ , y el régimen 2 con probabilidad $(1-\lambda)$. De esta forma, todos los retornos observados se consideran provenientes de una distribución mixta, y se estima por máxima verosimilitud el vector de parámetros $\theta = (\alpha_1, \beta_1, \sigma_1^2, \alpha_2, \beta_2, \sigma_2^2, \lambda)$. Esta especificación tiene la ventaja de que la probabilidad asociada con la densidad mixta implícitamente cubre todas las formas posibles de asignar observaciones a las dos regresiones, por las cuales los retornos podrían haber sido generados, y escoge la más probable. En su estudio hecho para los fondos mutuos de EE.UU. estos autores no encuentran evidencia para rechazar la hipótesis nula de estacionariedad en los parámetros, en el contexto de la línea de mercado de activos como en el CAPM de Black².

El modelo anterior sólo mide selectividad, ahora sin sesgo, pero no da una medida que permita detectar timing. Un año más tarde, en Kon y Jen (1979), proponen un modelo para medir tanto la habilidad de timing como la selectividad. El método es también de “regresión cambiante”, pero ahora se le añade una condición técnica de identificación. Lo ventajoso de este modelo, es que permite ser estimado en forma empírica en el contexto de la línea de mercado de activos y del CAPM de Black.

² Incidentalmente, Kon y Jen sugieren que este método contribuye a resolver un problema empírico frecuente. En efecto, diferentes autores han notado que el modelo CAPM discrimina en contra de los portafolios de alto riesgo y en favor de los portafolios de bajo riesgo [Friend y Blume (1970), Klemkosky (1973); McDonald (1974)]. A los primeros les exige demasiado, dejando a muchos portafolios riesgosos debajo de la línea de mercado en forma innecesaria; lo contrario ocurre para los portafolios de bajo riesgo, en forma no merecida quedan sobre la línea de mercado, pues se les exige demasiado poco. Kon y Jen atribuyen este problema al hecho de que el CAPM supone que los parámetros son estables, luego el método de mínimos cuadrados estima estos parámetros con sesgo.

5. Condiciones económicas cambiantes

Ferson y Schadt (1996) advierten sobre la importancia de incorporar condiciones económicas cambiantes al evaluar el desempeño de fondos mutuos. Los autores muestran que si los administradores cambian sus portafolios basados en información pública sobre las condiciones económicas, las medidas de desempeño tradicionales que no toman en cuenta este efecto, están sesgadas. Por ello, proponen la llamada “medida de Jensen condicionada”. En su modelo, el beta del portafolio administrado se supone función lineal del vector Z_t de información pública sobre las condiciones de la economía (que cambian en el tiempo):

$$\beta_j(Z_t) = \beta_{1j} + \beta_{2j}z_j$$

Donde β_{1j} es la media del beta condicional, $E(\beta_j(Z_t))$. El coeficiente β_{2j} representa la sensibilidad al vector de innovaciones en la información pública, $z_t = Z_t - E(Z_t)$.

Multiplicando el retorno del mercado por $(\beta_j(Z_t))$, se obtiene la siguiente ecuación:

$$r_{pt} - r_{lt} = \alpha_p + \beta_{1p}[r_{mt} - r_{lt}] + \beta_{2p}[z_{t-1}(r_{mt} - r_{lt})] + \varepsilon_{pt}$$

El factor adicional en la ecuación se puede interpretar como una estrategia dinámica autofinanciada que consiste en comprar z_{t-1} unidades del portafolio de mercado, financiada con deuda libre de riesgo.

B. Otras medidas de desempeño financiero

En esta sección se describen medidas de desempeño fuera del modelo de dos parámetros.

1. Medidas de desempeño basadas en momentos de orden superior

Una forma de derivar el modelo CAPM es suponer que los retornos siguen una distribución normal multivariada. En este caso, los dos primeros momentos estadísticos son suficientes para describir toda la distribución.

Sin embargo, estudios de retornos en Estados Unidos y en Chile indican que los retornos accionarios siguen más una distribución con riesgo variable, o, en su defecto, una distribución estable del tipo Pareto-Levi, Gregoire (1984). Esto sugiere que momentos de orden superior son relevantes para describir la distribución de los retornos.

Arditti (1977) encuentra evidencia que el tercer momento central (sesgo o asimetría) sería valorado por los inversionistas: al regresionar retornos medios contra desviación estándar y tercer momento, encuentra una relación negativa y significativa entre retorno medio y tercer momento (aunque la relación con desviación estándar es no significativa, consistente con CAPM). Es decir, los inversionistas estarían dispuestos a sacrificar retorno promedio para obtener mayor sesgo (colas más extendidas a la derecha, reflejando mayor probabilidad de retornos altos).

Ello le lleva a proponer la siguiente metodología para comparar el desempeño de portafolios: además de comparar según índices de Sharpe, considerar el sesgo normalizado por la medida de dispersión desviación estándar:

El portafolio i tiene mejor desempeño que el portafolio j , si se cumplen simultáneamente las siguientes condiciones:

$$S_i \geq S_j$$

$$\frac{M_{3i}}{\sigma_i} \geq \frac{M_{3j}}{\sigma_j}$$

El cociente M_{3i}/σ_i se obtiene de extraer la raíz cúbica del coeficiente de sesgo (*skewness*, que es la medida que entrega el programa TSP 70).

Por otra parte, si el coeficiente de Sharpe del portafolio i supera al del portafolio j , pero no el tercer momento, no se puede concluir que el portafolio i haya tenido un desempeño superior al j .

Arditti evaluó el desempeño de 34 fondos mutuos estudiados previamente por Sharpe (1966) en el período 1944-1963, con resultados que anulan los encontrados por Sharpe: Sharpe encontró que los fondos mutuos habían tenido un desempeño inferior al índice Dow-Jones, pero Arditti indica que ello se compensa por sesgo superior a dicho índice.

En principio, se podría pensar que también los momentos cuarto, quinto, etc. sean de utilidad para describir la distribución de los retornos. Ésta es la visión que toman los criterios de dominancia estocástica, que se basan en la función de probabilidades completa, con supuestos poco restrictivos sobre las preferencias de los inversionistas (por ejemplo, el criterio

de dominancia estocástica de primer orden supone que los inversionistas prefieren más a menos, y el de segundo orden supone que los inversionistas son aversos al riesgo, etc.). Esta metodología tiene el inconveniente que no da un orden completo, esto es, no permite comparar todos los portafolios (no permite concluir si un portafolio tuvo mejor desempeño que otro, o si fueron iguales). Joy y Porter (1974) estudiaron los mismos fondos mutuos que Sharpe y Arditti, utilizando el criterio de dominancia estocástica. De los treinta y cuatro fondos comparados con el índice Dow Jones, la dominancia estocástica de primer orden no permitió la comparación de ningún portafolio, la de segundo orden permitió seis comparaciones, y la de tercer orden permitió nueve. Sin embargo, en todos los casos se encontró que los fondos tuvieron un desempeño inferior al del índice, lo que tendería a confirmar el resultado inicial de Sharpe.

2. Medidas basadas en teoría de opciones

Hendrikson y Merton (1981) proponen un test no paramétrico para testear evidencia de timing y selectividad, el cual no usa el CAPM como base, ni requiere distribución normal por parte de los retornos de los portafolios. Los autores confeccionan un modelo basado en teoría de opciones.

El problema con este modelo es que requiere que los pronósticos de los inversionistas respecto a movimientos del mercado sean observables, lo cual es una limitante fuerte para pruebas empíricas. A esto se suma la ya mencionada deficiencia, de que sólo entrega signos y no magnitudes.

3. Medidas basadas en el modelo de precios por arbitraje (APT)

Otro modelo de valoración de activos ampliamente tratado en la literatura es el modelo APT. Suponiendo que los retornos son generados por k factores de riesgo, se puede mostrar que en ausencia de oportunidades de arbitraje el retorno esperado del activo o portafolio i se relaciona linealmente con las sensibilidades del activo con respecto a los k factores de riesgo:

$$m_i = r_f + b_{i1} \lambda_1 + \dots + b_{ik} \lambda_k$$

donde

m_i = retorno esperado del activo i

r_f = tasa libre de riesgo

b_{ij} = sensibilidad del retorno del activo i frente al factor j .

λ_j = premio por riesgo asociado al factor j

Este modelo permite definir una medida de desempeño análoga a la medida de Jensen de la siguiente forma:

$$m_i = r_I + a_i + b_{iI} \lambda_I + \dots + b_{ik} \lambda_k$$

donde a_i es una medida del retorno anormal del activo i .

4. Modelo de Grinblatt y Titman (1989)

Este modelo no requiere de la identificación de un portafolio de mercado, ni se basa en el CAPM ni en el APT. El modelo supone que hay inversionistas que están mejor informados que otros sobre movimientos futuros del mercado (*timing*) y características individuales de activos (selectividad), y también que las señales de *timing* y selectividad son independientes. Otro supuesto es que la distribución de los retornos es normal multivariada, con parámetros poblacionales estables en el tiempo.

La idea es prescindir del portafolio de mercado, usando como portafolio de referencia algún portafolio eficiente en el conjunto de los activos transables. Por ser eficiente *ex post*, cualquier activo transable no debiera obtener mejor desempeño que este portafolio de referencia (Roll 1978).

Para cada activo individual, se estima la sensibilidad de este activo con respecto al portafolio eficiente seleccionado, denotada b_i , aunque no corresponda al riesgo de mercado. El beta del portafolio sigue siendo el promedio ponderado de los betas de los activos que lo componen. Si se trabaja con excesos de retornos (retornos por sobre la tasa libre de riesgo), se puede establecer la siguiente relación matemática:

$$r_{pt} - r_{It} = b_{pt-I}(r_{et} - r_{It}) + e_{pt}$$

Como mostró Roll, el coeficiente alfa debe ser cero y por ello no aparece en la ecuación. Tomando esperanza matemática al exceso de retorno, obtenemos que:

$$E(r_{pt} - r_{It}) = E(b_{pt-I})E(r_{et} - r_{It}) + \text{cov}(b_{pt-I}, r_{et} - r_{It}) + E(e_{pt})$$

Si alguno de los dos últimos términos de la ecuación es distinto de cero, ello indica retorno anormal. Suponiendo que el beta de cada activo individual es estable en el período de evaluación, los dos últimos términos se pueden atribuir a una administración activa de portafolio. Un valor positivo de la covarianza refleja *timing*, y un valor positivo de la esperanza del error, selectividad.

II. DISCUSIÓN

En esta sección discutimos la elección de una medida de desempeño ajustada por riesgo que sea apropiada para la industria de pensiones en Chile.

A. *Benchmark*

Como se ha señalado, un *benchmark* debe ser relevante (posible de elegir por el inversionista) y alcanzable (debe tener restricciones similares a las del portafolio que se evalúa). En el caso de los fondos de pensiones chilenos, el primer requisito deja fuera a cualquier portafolio distinto al administrado por una AFP, porque estas sociedades de giro exclusivo son las únicas que pueden administrar fondos previsionales. Similarmente, el segundo criterio impide utilizar cualquiera de los índices bursátiles tradicionales (IGPA, IPSA, GLOBAL) porque (i) los FP se invierten en bonos además de acciones y no existen índices de bonos ni índices mixtos, y (ii) los FP enfrentan restricciones de portafolio, las que han ido variando en el tiempo y que además dependen del tamaño de cada fondo. Por este motivo, se propone comparar el desempeño de los fondos entre sí, y no contra algún índice de mercado, y entender las comparaciones desde el punto de vista de un afiliado, que desea saber a qué fondo le fue mejor al incluir el riesgo que tomó.

Conviene destacar que debido a restricciones de portafolio diferentes, la presente investigación no permite sacar conclusiones inambiguas respecto de la calidad de la administración de cartera de las administradoras. En efecto, tanto Walker (1993) como Arrau y Chumacero (1998)³ encuentran que existe una relación sistemática entre los retornos de los fondos de pensiones y el tamaño de los mismos, la que puede ser atribuida (además de diferentes habilidades de administración), a restricciones dependientes del tamaño de los fondos⁴. En este sentido, las mediciones de desempeño que nosotros proponemos no permiten comparar inambiguamente la calidad profesional de los equipos de inversión de las AFPs, pero sí permiten mostrar a los afiliados qué AFPs obtuvieron mejores resultados, dado el riesgo de los portafolios en que invirtieron.

³ Los autores no encuentran la misma relación, sin embargo, Walker encuentra una relación entre rentabilidad y tamaño en los fondos de renta variable, separando los fondos en grandes y pequeños. Arrau y Chumacero concluyen que la relación no es lineal, lo que se explicaría por las restricciones más fuertes que enfrentan los fondos grandes, y por los incentivos de los pequeños a seguir las inversiones de los grandes, debido a la regulación de rentabilidad mínima que penaliza desviaciones (negativas) de la rentabilidad promedio. Por ello, los fondos de tamaño medio alcanzarían retornos mejores que los grandes y que los chicos.

⁴ Por ejemplo, las restricciones por emisor asociadas a la participación accionaria.

B. Riesgo total versus riesgo sistemático

Como se ha señalado, el concepto de riesgo total de los fondos previsionales es aplicable a algunos cotizantes, y el concepto de riesgo de mercado a otros. Generalmente al momento de jubilar el capital humano tiene un valor económico pequeño, por lo que la diferencia depende de qué otros activos tengan los trabajadores al jubilar. Si algunos tienen un porcentaje importante de su riqueza invertido en acciones, entonces el coeficiente beta del portafolio de las AFPs (calculada con respecto a algún índice de mercado como el IGPA o el GLOBAL) es la medida apropiada. Nuestro argumento es que el concepto de riesgo total es preferible en este caso, porque (i) a la fecha de jubilación los fondos previsionales presumiblemente representan un alto porcentaje de la riqueza de la mayoría de los cotizantes, (ii) relativamente pocos individuos mantienen parte importante de su riqueza en acciones en Chile, y (iii) no existen índices para otras formas de riqueza como bonos, y propiedades; por lo que no es posible calcular covarianzas con respecto a ellos.

C. Modelo de retornos de activos

Las medidas de desempeño basadas en el modelo CAPM están sujetas a varias críticas, principalmente: (i) requieren identificar el portafolio de mercado; pero como éste es inobservable se utiliza un proxy. El problema es que la matemática del conjunto eficiente implica que pequeños cambios en el índice elegido podrían alterar el orden de portafolios sobre la base de desempeño (Roll 77), (ii) bajo información asimétrica, y con un índice de mercado ineficiente, el análisis de línea de mercado de activos puede llevar a errores significativos porque tanto portafolios eficientes como ineficientes pueden ubicarse fuera de la línea de mercado de activos, Dybig y Ross, 1985⁵.

Además, algunos artículos estudian empíricamente si los ordenamientos de desempeño financiero obtenidos usando como base el CAPM y el APT difieren mucho. Por ejemplo, Lehmann y Modest (1987) encuentran que los ordenamientos difieren bastante, dependiendo de qué modelo de valoración se use. Peor aún, obtienen que los ordenamientos también dependen de la especificación particular que se haga del APT, lo que

⁵ Afortunadamente, estos problemas no parecen ser importantes empíricamente. Por ejemplo, Peterson y Rice (1980) utilizan diferentes índices de mercado para evaluar el desempeño de administradoras en Estados Unidos, obteniendo órdenes (ranking) bastante parecidos.

enfatisa la necesidad de identificar el verdadero modelo de valoración de activos.

Debido a estos problemas, es clara la conveniencia de utilizar una medida de desempeño ajustada por riesgo que no requiera especificar un modelo de valoración de activos.

D. Condiciones económicas cambiantes y timing del mercado

Dada la poca profundidad del mercado de capitales chileno, y el tamaño de los fondos de pensiones que limita cambios en los portafolios ante información pública sobre variables de la economía, estimamos que la medida de Ferson y Schadt para incluir posible respuesta a condiciones económicas es menos relevante en nuestro estudio de fondos de pensiones, que lo que sería en un estudio del desempeño de fondos mutuos. Por los mismos motivos, y en vista de los resultados encontrados por Walker, consideramos que los fondos no pueden ser *timers* activos. Por ello, no estimamos necesario incluir medidas que incluyan el *timing* del mercado.

E. Proposición de una medida de desempeño financiero para los fondos de pensiones en Chile

Sobre la base de la discusión anterior, se propone utilizar la medida de desempeño de Sharpe para rankear los fondos de pensiones, y siguiendo la práctica financiera, se propone realizar los cálculos sobre períodos de 60 meses. Esta medida presenta las siguientes ventajas: (i) ocupa el concepto de riesgo total, más apropiado para fondos de pensiones si (como se ha argumentado), éstos constituyen la mayor parte de la riqueza de los pensionados a la fecha de jubilación, (ii) no depende de un modelo específico de valoración de activos de capital, y (iii) por lo mismo, no requiere identificar el portafolio de mercado. No es posible ordenar los fondos utilizando un criterio que incluya más de dos momentos (Arditi, dominancia estocástica), porque no dan un orden completo. Sin embargo, se propone reportar adicionalmente el coeficiente de Arditi para incluir el sesgo, aparentemente valorado por los inversionistas.

La medición propuesta, sin embargo, presenta como limitación el hecho de no controlar por las diferencias en liquidez de las inversiones. En

la actual regulación, las diferencias de liquidez implican riesgos de valorización, porque la cartera completa puede ser valorizada por el precio de mercado de una transacción de tamaño pequeño. Así, aunque el precio en una transacción de poco monto no es tan buen estimador del valor económico de un activo, ese precio es utilizado para valorizar la cartera de los fondos, y afecta la rentabilidad de la cuota⁶.

III. DATOS Y EVALUACIÓN DE DESEMPEÑO FINANCIERO

A. La muestra

La muestra consiste de los retornos mensuales reales (deflactados por UF) de las cuotas de los nueve fondos de pensiones existentes entre enero de 1987 y septiembre de 1998: Cuprum, Habitat, Magister, Plan Vital, Protección, Provida, Santa María, Summa y también se incluye Unión a pesar de que ésta desaparece en junio de 1998 al fusionarse con Provida permaneciendo esta última. Están incluidas pero bajo otro nombre debido a fusión Invierta, El Libertador, Banguardia, Concordia y Futuro⁷. No fueron incluidos, porque fueron creadas con posterioridad a 1987, Valora (nace en agosto de 1990 con el nombre de Norprevisión), Armoniza (nace en agosto de 1996), Qualitas (nace en diciembre de 1996), Previpán, Bansander (nace en agosto de 1996), y Genera (nació en mayo de 1993 y se liquidó en agosto de 1995). De este modo, nuestros resultados no parecen estar afectados por sesgo de sobrevivencia, al incluir todos los fondos que existían en enero de 1987.

⁶ Una solución posible es rankear las AFPs por clases de activos. Sin embargo, ello tiene la desventaja (respecto del método propuesto) que al inversionista le interesa también qué porcentaje invierten en cada clase de activo. Por otro lado, este problema está siendo abordado por la autoridad, que estudia la posibilidad de cambiar el método de valorización de algunas inversiones: actualmente toda la cartera se valora a liquidación de mercado, como si fuese vendida cada día, y se está considerando incorporar un método diferente para activos que las AFP se propongan mantener a largo plazo.

⁷ Las fusiones fueron: Plan Vital e Invierta (1 de diciembre de 1993, siguiendo como Plan Vital), Provida y El Libertador (1 de mayo de 1995, siguiendo como Provida), Santa María y Banguardia (1 de mayo de 1995, siguiendo como Santa María), Valora y Qualitas (1 de junio de 1995, siguiendo como Valora), Valora y Previpán (1 de noviembre de 1995, siguiendo como Valora), Plan Vital y Concordia (1 de marzo de 1996, siguiendo como Plan Vital), Armoniza y Valora (1 de noviembre de 1996, siguiendo como Qualitas), Magister y Futuro (permaneciendo Magister).

B. Ranking de los fondos de pensiones

La Tabla N° 1 contiene indicadores de desempeño de los nueve fondos en el período completo⁸. Según ella, la rentabilidad real promedio (deflactada por unidad de fomento y anualizada) fluctuó entre 7,32% (Provida) y 9,21% (Unión), y la desviación estándar de la rentabilidad fluctuó entre 6,11% (Provida) y 6,81% (Magister). Es decir, la AFP de mínima rentabilidad fue también la de menor riesgo en el período. Finalmente, en el período completo los órdenes según rentabilidad promedio y el índice de Sharpe fueron muy similares, obteniéndose un coeficiente de correlación de Spearman de 0,96. Como veremos más adelante, la correlación entre rankings se deteriora al analizar subperíodos.

Además, no es posible rechazar la hipótesis nula de que los retornos siguen una distribución normal. Para todos los fondos, el índice de Jarque-Bera es menor al límite que entregan las tablas de chi-cuadrado con 5% de significancia y dos grados de libertad (5.99), lo cual permite aceptar la hipótesis nula de Normalidad.

Al aplicar el criterio de Arditti, encontramos que es posible comparar de a pares las AFPs entre sí sólo en el 30,5% de los casos⁹.

Las Tablas N°s. 2 y 3 reportan el desempeño relativo de los fondos de pensiones según el criterio de Sharpe, en períodos (sobrepuestos entre sí) de 3 y 5 años

De las Tablas N°s. 2 y 3 se desprende que el tamaño de la muestra utilizado para efectuar los rankings afecta el resultado en forma importante.

⁸ En este trabajo se examinó el desempeño relativo de las cuotas (incluyendo renta variable y renta fija). Ello implica un error al estimar las desviaciones estándar de los retornos, porque la desviación estándar de los retornos del componente renta fija no es constante. Específicamente, la desviación estándar de un bono disminuye al pasar el tiempo porque los bonos de más largo plazo económico (duración) son más sensibles a los movimientos en la tasa de interés, con lo cual no sería estable en el período de estimación. Por supuesto, en el caso de un portafolio, ello depende de la duración de los nuevos títulos de renta fija que son comprados, y este error es pequeño en la medida en que el plazo económico de los portafolios de renta variable de las AFPs sean estables en el período de estimación. La Superintendencia reporta los plazos económicos de las carteras sólo a partir de enero de 1991, y en ese período ellos parecen relativamente estables: el promedio de cambios porcentuales de plazo económico es inferior a 0,2%, con desviaciones estándar que no superan el 4% en general. Además, en períodos mensuales de 3 años ninguna cartera de renta fija aumenta ni disminuye sistemáticamente los plazos económicos de sus inversiones todos los meses. Por ello, el error que se comete al aplicar estos indicadores a la cartera completa (y no sólo a la parte renta variable de los portafolios) no parece importante.

⁹ De 36 pares, sólo fue posible comparar 11: Cuprum tuvo mejor desempeño que Plan Vital; Habitat peor que Protección, pero mejor que Provida y que Santa María; Magister tuvo mejor desempeño que Provida y Santa María; Plan Vital tuvo peor desempeño que Protección; ésta tuvo mejor desempeño que Provida, Santa María y Suma, y Provida tuvo peor desempeño que Santa María.

TABLA N° 1: DESEMPEÑO FINANCIERO DE LOS FONDOS DE PENSIONES: ENERO 1987 - SEPTIEMBRE 1998

	Cuprum	Habitat	Magister	Planvital	Protección	Provida	Sta. María	Summa	Unión
Promedio	8,550	7,946	7,984	8,276	8,555	7,324	7,560	8,109	9,216
D. estandard	6,717	6,371	6,881	6,696	6,752	6,112	6,448	6,593	6,519
Máximo	71,558	72,648	81,744	76,389	77,797	71,644	69,925	72,015	72,960
Mínimo	-65,104	-55,302	-58,197	-56,302	-59,401	-50,201	-56,198	-59,962	-45,470
Sego	0,2572	0,2804	0,3480	0,2430	0,2821	0,2693	0,2705	0,2576	0,2042
Curtosis	3,1675	3,2917	3,4889	3,2990	3,4061	3,3978	3,2689	3,2118	3,2118
Jarque-Bera	1,6022	2,1953	3,989	1,7913	2,6664	2,4734	2,3974	1,8563	1,163
Probability	0,4488	0,3336	0,136	0,4079	0,2636	0,2903	0,3015	0,3952	0,559
Sharpe	0,3608	0,2856	0,2699	0,3210	0,3597	0,1959	0,2223	0,03007	0,4739
Arditti	0,6359	0,6545	0,7034	0,6241	0,6559	0,6457	0,6468	0,6363	0,5889
Ranking Prom.	3	7	6	4	2	9	8	5	1
Ranking sharpe	2	6	7	4	3	9	8	5	1

TABLA N° 2: RANKING DE LOS FONDOS DE PENSIONES SEGÚN SHARPE
(Períodos de tres años)

	Cuprum	Habitat	Magister	Planvital	Protección	Provida	Sia. María	Summa	Unión
Ene87-Dic89	1	7	6	4	3	9	8	5	2
Ene88-Dic90	3	7	6	1	2	9	8	5	4
Ene89-Dic91	2	6	7	3	4	9	8	1	5
Ene90-Dic92	3	4	6	7	2	9	8	1	5
Ene91-Dic93	4	3	6	9	2	8	5	1	7
Ene92-Dic94	2	6	9	4	1	3	8	5	7
Ene93-Dic95	1	6	9	3	2	7	8	4	5
Ene94-Dic96	1	5	7	2	3	6	9	8	4
Ene95-Dic97	7	1	9	5	4	6	8	2	3
Oct95-Sep98	9	2	7	4	8	6	5	3	1

TABLA N° 3: RANKING DE LOS FONDOS DE PENSIONES SEGÚN SHARPE
(Períodos de cinco años)

	Cuprum	Habitat	Magister	Planvital	Protección	Provida	Sia. María	Summa	Unión
Ene89-Dic91	1	6	7	3	4	9	8	2	5
Ene90-Dic92	1	7	6	4	2	9	8	3	5
Ene91-Dic93	3	6	7	5	2	9	8	1	4
Ene92-Dic94	2	5	7	4	1	9	8	3	6
Ene93-Dic95	2	5	4	7	1	9	8	3	6
Ene94-Dic96	2	6	8	3	1	7	9	5	4
Ene95-Dic97	2	4	9	3	1	7	8	6	5
Oct95-Sep98	6	3	9	2	5	4	7	8	1

Por ejemplo, Summa aparece rankeada 3 en el último período si toma 3 años, pero 8 si toma 5, y Cuprum aparece en lugar 9 tomando 3 años pero 6 tomando 5 años, en el mismo período. En este sentido, la práctica interna-cional es considerar períodos de 5 años para evaluación.

La Tabla N° 4 reporta los coeficientes de correlación entre los orde-namientos según desempeño Promedio y Sharpe, considerando períodos de 36 meses (normativa actual) y 60 meses (período de evaluación propuesto). En la Tabla N° 4 se aprecia que las correlaciones entre los rankings son mayores mientras más largo el intervalo de evaluación. Por ejemplo, la correlación promedio entre ambos órdenes en períodos de tres años fue 0,81 mientras que en períodos de cinco años alcanzó a 0,91 (ambos inferior-es a la correlación en el período completo). Pero conviene notar que a pesar de la alta correlación promedio en períodos de tres años, ella fluctúa entre 0,45 (enero 92-diciembre 94) y 0,98 (enero 94-diciembre 96), por lo cual los dos rankings no pueden considerarse equivalentes. Por otra parte, en períodos de 5 años las correlaciones entre rankings promedio y corregi-dos por riesgo son consistentemente superiores a 0,85, por lo cual podrían considerarse similares.

TABLA N° 4: CORRELACIÓN ENTRE ORDENAMIENTOS SEGÚN PROMEDIO Y SHARPE

Evaluación en períodos de 3 años		Evaluación en períodos de 5 años	
Ene87-Dic89	0,9667	Ene87-Dic91	0,8500
Ene88-Dic90	0,9500	Ene88-Dic92	0,8667
Ene89-Dic91	0,7667	Ene89-Dic93	0,8500
Ene90-Dic92	0,8167	Ene90-Dic94	0,8833
Ene91-Dic93	0,6333	Ene91-Dic95	0,9333
Ene92-Dic94	0,4500	Ene92-Dic96	1,0000
Ene93-Dic95	0,9333	Ene93-Dic97	0,9500
Ene94-Dic96	0,9833	Oct93-Sep98	0,9833
Ene95-Dic97	0,7500		
Oct95-Sep98	0,8500		
Promedio	0,8100	Promedio	0,9146

C. Persistencia del retorno de los fondos

Inmediatamente que los resultados son conocidos, es usual que mu-chos vendedores salgan a captar nuevos afiliados mostrando los resultados reportados por la Superintendencia. ¿Es racional para un cotizante cambia-rse de AFP basado en el desempeño histórico? Ello depende (entre otros

factores) de si los resultados exhiben persistencia o no. En efecto, si la correlación entre los ordenes fuese negativa, sería irracional: los cotizantes que siguieran el consejo de dichos vendedores perderían dinero (entrarían caro y saldrían barato). Para examinar esta pregunta, en las Tablas N^{os} 5 y 6 reportamos los coeficientes de correlación en el tiempo de ranking de desempeño según el retorno promedio y según Sharpe. En ellas, la columna 1 reporta los ocho coeficientes de correlación serial entre los nueve subperíodos consecutivos (1-2, 2-3, 3-4, 4-5, 5-6, 6-7, 7-8 y 8-9), la columna 2 reporta los siete coeficientes de correlación serial entre los subperíodos 1-3, 2-4, 3-5, 4-6, 5-7, 6-8 y 7-9, y así sucesivamente.

Debido a que la columna 1 implica comparar períodos en que se superponen 2 de los 3 años considerados, no es extraño encontrar que en muchos de ellos la correlación temporal es significativamente distinta de cero con 1% de significancia: simplemente se está considerando el período consigo mismo en un 67%. En la columna 2 consideramos períodos que sólo tienen 1 año en común, y se observa que sólo en dos subperíodos (medido por Sharpe) la correlación temporal fue positiva (al 5%), siendo estadísticamente no distinta de cero en los demás.

En la Tabla N^o 5 (desempeño promedio sin ajustar por riesgo) se aprecia que si bien en 5 períodos aparecen correlaciones entre períodos no superpuestos positivas al 5%, en el último período aparecen correlaciones negativas estadísticamente distintas de cero al 5%. Ello indica que en este último período no fue conveniente (expost) elegir la AFP por el desempeño de los últimos tres años.

En la tercera columna (primera en que no hay traslape alguno) de la Tabla N^o 6 (desempeño ajustado por riesgo), sólo en un subperíodo la correlación fue positiva con nivel de significancia 5%, y en los demás fue cero estadísticamente (incluso al final de muestra fue negativa aunque no significativa). Además, se observa en general que las correlaciones promedio van disminuyendo a medida en que los subperíodos están más alejados entre sí. La conclusión general es que hay alguna evidencia que apoya la racionalidad de cambiarse en base a desempeño, pero que ella es más bien débil.

Las Tablas N^{os} 7 y 8 reportan el mismo ejercicio para mediciones de promedio y ajustadas por riesgo (Sharpe) sobre períodos de evaluación de cinco años. Las conclusiones son similares, aunque (como era de esperarse por la superposición de períodos) se observan correlaciones más altas al considerar períodos de evaluación más largos.

Finalmente, conviene notar que en esta sección nos hemos centrado exclusivamente en la rentabilidad. Evidentemente, para tomar una decisión

TABLA N° 5: CORRELACIÓN TEMPORAL DE ÓRDENES SEGÚN PROMEDIO
(Períodos de tres años)

	1	2	3	4	5	6	7	8	9
0,9333*									
0,7833*		0,8667*							
0,9167*		0,6000**	0,6833**						
0,9167*		0,8167*	0,3833	0,5667					
0,3333		0,3667	0,5500	0,7333**	0,8167*				
0,6000**		0,0000	0,3000	0,5500	0,8667*	0,7167**			
0,8167*		0,7833*	0,1000	0,2833	0,5500	0,7833*	0,8000*		
0,5667		0,5500	0,1167	0,1667	0,3167	0,5000	0,4000	0,4000	
-0,0500		-0,2333	-0,1500	-0,4500	-0,6833**	-0,6333**	-0,6167**	-0,4500	-0,4833
Promedio	0,6463	0,4688	0,2833	0,3083	0,3733	0,3417	0,1944	-0,0250	-0,4833

(*) Significativo al 1%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,783.

(**) Significativo al 5%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,600.

(***) Significativo al 10%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,572.

TABLA N° 6: CORRELACIÓN TEMPORAL DE ÓRDENES SEGÚN SHARPE
(Períodos de tres años)

	1	2	3	4	5	6	7	8	9
0,8500*									
0,7667**		0,7500**							
0,7833*		0,4833	0,6000**						
0,8333*		0,4333	0,0333	0,1500					
0,1833		0,3333	0,4167	0,4500	0,5667				
0,8000*		0,2833	0,6167	0,8167*	0,8000*	0,7833*			
0,7833*		0,6500**	-0,1833	0,2333	0,4667	0,7500**	0,7333**		
0,1333		0,3500	0,2167	0,4000	0,5353	0,4000	0,1667	0,1500	
0,6667**		-0,2667	0,2333	-0,4667	-0,1000	-0,0333	0,0000	-0,1500	-0,1500
Promedio	0,6444	0,3771	0,2095	0,2639	0,4133	0,4750	0,3000	0,0000	-0,1500

(*) Significativo al 1%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,783

(**) Significativo al 5%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,600

(***) Significativo al 10%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,572

TABLA N° 7: CORRELACIÓN TEMPORAL DE ÓRDENES SEGÚN RETORNO PROMEDIO
(Períodos de cinco años)

	1	2	3	4	5	6	7
	0,8667*						
	0,9833*	0,9000*					
	0,9833*	0,9667*	0,8667*				
	0,9667*	0,9833*	0,9500*	0,8500*			
	0,6500**	0,7167**	0,7000**	0,6500**	0,8500*		
	0,8833*	0,3333	0,4500	0,3833	0,3333	0,5833***	
	0,5333	0,3333	-0,3500	-0,3500	-0,3167	-0,3333	0,0167
Promedio	0,8381	0,7056	0,5233	0,3833	0,2889	0,1250	0,0167

(*) Significativo al 1%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,783

(**) Significativo al 5%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,600

(***) Significativo al 10%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,572

TABLA N° 8: CORRELACIÓN TEMPORAL DE ORDENES SEGÚN SHARPE
(Períodos de cinco años)

	1	2	3	4	5	6	7
	0,9333*						
	0,9000*	0,8833*					
	0,9000*	0,9333*	0,8833*				
	0,8500*	0,8000*	0,8333*	0,6833**			
	0,6167**	0,8667*	0,7667*	0,8500*	0,7833*		
	0,9333*	0,5333	0,8333*	0,6333**	0,7167**	0,6833**	
	0,4667	0,3833	-0,3500	0,0000	-0,3333	0,0833	0,0167
Promedio	0,8000	0,7333	0,5933	0,5417	0,4889	0,3000	0,0167

(*) Significativo al 1%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,783

(**) Significativo al 5%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,600

(***) Significativo al 10%, para testear la hipótesis nula coeficiente de correlación de Spearman = 0 contra coeficiente > 0. El valor crítico para 9 observaciones es 0,572

racional de cambiarse o no de AFP importa también la comisión de la administradora. Este elemento de costo es particularmente importante en los primeros años de cotización, puesto que la rentabilidad se aplica sobre un fondo acumulado menor. En otras palabras, a lo largo de la vida de un cotizante, podría ser racional elegir al inicio la AFP con menor comisión absoluta, y a medida que el fondo acumulado aumente, buscar la que le entregue una mayor rentabilidad. Nuestro aporte en esta sección ha consistido en ver si el desempeño histórico es de utilidad para estimar el desempeño futuro del fondo, y luego es de mayor utilidad para cotizantes con varios años en el sistema.

IV. RESUMEN Y CONCLUSIONES

En este artículo revisamos la literatura de medición de desempeño financiero de portafolios, y discutimos su aplicabilidad a la industria de fondo de pensiones chilena. Se propuso medir los fondos según el índice de Sharpe, porque (i) se basa en el concepto de riesgo total, más apropiado para la mayoría de los pensionados a la fecha de jubilación, (ii) no depende de un modelo específico de valoración de activos, y (iii) puesto que en particular no se basa en el CAPM, no requiere identificar el portafolio de mercado.

Posteriormente se aplicó este criterio para realizar un ranking de los fondos de pensiones, el que fue comparado con el orden de desempeño según promedio. Este análisis entregó dos conclusiones principales:

Utilizando períodos de 3 años para medir, no es indiferente qué medida utilizar. En efecto, el coeficiente de correlación (Spearman) entre rankings basados en rentabilidad promedio (criterio actual) y ajustados por riesgo (Sharpe) fluctúa entre 0,45 y 0,98, por lo que en algunos períodos los órdenes difieren bastante. Esta diferencia se atenúa si se consideran períodos de cinco años, en cuyo caso las correlaciones no fueron inferiores a 0,85 en ningún período.

Encontramos alguna evidencia de la racionalidad de cambiarse de AFP conociendo el desempeño histórico, pero no mucha. En efecto, al comparar los desempeños (ajustados por riesgo) entre períodos no superpuestos, se observa que en general las correlaciones seriales son positivas (aunque sólo en un 20% de los períodos estadísticamente distintas de cero).

REFERENCIAS BIBLIOGRÁFICAS

- Ang, S.; y Chua, J. H. "Composite measures for the evaluation of investment performance". *Journal of Financial and Quantitative Analysis*, junio 1979.
- Arditt, D. "Another look at mutual fund performance". *Journal of Financial and Quantitative Analysis*, junio 1971.
- Arrau, P.; y Chumacero, R. "Tamaño de los fondos de pensiones en Chile y su desempeño financiero". *Cuadernos de Economía*, 35 N° 105 (agosto 1998), pp. 205-235.
- Black, M.C. Jensen; y Scholes, M. C. "The capital asset pricing model: Some empirical tests". En M. C. Jensen (editor), *Studies in the theory of capital markets*. Nueva York: Praeger, 1972.
- Cai, Jun; K.; Chan, C.; y Yamada, Takeshi. "The performance of Japanese Mutual Funds". *The Review of Financial Studies*. Vol. 10, N° 2 (verano 1997), pp. 237-273.
- Coggin, Daniel; Fabozzi, Frank J.; y Rahaman, Shafiqur. "The investment performance of US equity pension fund managers: An empirical investigation". *Journal of Finance* 48, N° 3 (julio 1993), pp. 1039-1055.
- Copleand; y Myers, D. "The value line enigma (1965-1978): A case study of performance evaluation issues". *Journal of Financial Economics* N° 10, 1982, pp. 289-321.
- Cumby; y Modest, D. "Testing for market timing ability". *Journal of Financial Economics* 19, 1987, pp. 169-189.
- Diets, Peter. O. "Components of a measurement model, rate of return, risk and timing". *Journal of Finance*, mayo 1968.
- Dybig; y Ross, S. "The analytics of performance measurement using a security market line". *Journal of Finance* 40, N° 2 (junio 1985).
- Fama, Eugene F. "Components of investment performance". *Journal of Finance* 27, N° 3 (junio 1972), pp. 551-567.
- Ferson, E; y Schadt, R. W. "Measuring fund strategy and performance in changing economic conditions". *Journal of Finance*, Vol. 51, 1996, pp. 425-461.
- Friend; y Blume, M. "Measurement of portfolio performance under uncertainty". *American Economic Review*, septiembre 1970.
- Gregoire. "Distribuciones Pareto-Levy para distribuciones de retornos en Chile". En J. Gregoire y S. Zurita, *Lecturas de Economía Financiera*. Editorial FACEA, Universidad de Chile, 1995.
- Grinblatt; y Titman, S. "Portfolio performance evaluation: old issues and new insights". *Review of Financial Studies* N° 2, pp. 393-421.
- . "Mutual fund performance: An analysis of quarterly portfolio holdings". *Journal of Business*, Vol. 62, N° 3.
- . "The persistence of mutual fund performance". *Journal of Finance*, Vol. 47, N° 5 (diciembre de 1992).
- Hendrikson, D.; y Merton, R. C. "On market timing and investment performance II. Statistical procedures for evaluating forecasting skills". *Journal of Business*, octubre 1981.
- Hendricks, Darryl; Patel, Jayendu; y Zekhauser, Richard. "Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988". *Journal of Finance*, Vol. 48, N° 1 (marzo 1993), pp. 93-130.
- Iglesias. "Estrategias de inversión de los fondos de pensiones". *Anales del Tercer Congreso de Finanzas*. Universidad Católica de Chile, agosto 1990.

- Jensen, Michael C. "The performance of mutual funds in the period 1945-1964". *Journal of Finance*, Vol. 23, N° 1 (mayo, 1968), pp. 167-247.
- . "Risk, the pricing of capital assets, and the evaluation of investment portfolios". *Journal of Finance*, abril 1969.
- Jobson, y Korkie, B. "Performance hypothesis testing with the Sharpe and Treynor Measure". *Journal of Finance*, Vol. 36, N° 4 (septiembre 1981).
- Joy, O.; y Porter, R. B. "Stochastic dominance and mutual fund performance". *Journal of Financial and Quantitative Analysis*, January 1974.
- Jon, Kon; y Jen, F. C. "Estimation of time-varying systematic risk and performance for mutual fund portfolios: An application of switching regression". *Journal of Finance*, mayo 1978.
- . "The investment performance of mutual funds: An empirical investigation of timing, selectivity and market efficiency". *Journal of Business*, abril 1979.
- Lehmann, Bruce N.; y Modest, David M. "Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons". *Journal of Finance*, junio 1987, pp. 233-265.
- Merton, C. "On market timing and investment performance I. An equilibrium theory of value for market forecasts". *Journal of Business*, julio 1981.
- Peterson, David; y Rice, Michael L. "A note on ambiguity in portfolio performance measures". *Journal of Finance*, Vol. 35, N° 5 (diciembre 1980), pp. 1251-1256.
- Roll, Richard. "A critique of the asset pricing theory's tests: Part I: On past and potential testability of the theory". *Journal of Financial Economics*, 7, 1977, pp. 391-400.
- . "Ambiguity when performance is measured by the security market line". *Journal of Finance*, Vol. 33, N° 4 (septiembre 1978), pp. 1051-1068.
- Sharpe, William. "Capital asset prices: A theory of market equilibrium under conditions of risk". *Journal of Finance*, 19 (septiembre 1969), pp. 425-442.
- . "Mutual fund performance". *Journal of Business*, Vol. 39, N° 1 (enero 1966), pp. 119-138.
- Schlarbaum, Gary G. "The investment performance of the common stock portfolios of property-liability insurance companies". *Journal of Financial and Quantitative Analysis*, enero 1974.
- Treynor, Jack L. "How to rate management of investment funds". *Harvard Business Review*, Vol. 43, N° 1 (enero-febrero 1965), pp. 63-75.
- Treynor, L.; y Mazuy, K. "Can mutual funds outguess the market?". *Harvard Business Review*, julio-agosto 1966.
- Walker, Eduardo. "Desempeño financiero de las carteras de renta fija de los fondos de pensiones en Chile. ¿Ha tenido desventajas ser grandes?". *Cuadernos de Economía*, 30, N° 89 (abril 1993), pp. 1-33.
- . "Desempeño financiero de las carteras accionarias de los fondos de pensiones en Chile. ¿Ha tenido desventajas ser grandes?". *Cuadernos de Economía*, 30, N° 89 (abril 1993), pp. 35-75.
- Wagner, Wayne H.; y Tito, Dennis A. "Definite new measures of bond performance and risk". *Pension World*, Vol. 13, N° 5 (mayo 1977), pp. 10-12. □