Banco de México

Documentos de Investigación

Banco de México

Working Papers

N° 2008-02

Analysis of the Performance of Mexican Pension Funds: Evidence from a Stationary Bootstrap Application

Arnulfo Rodríguez
Banco de México

Gerardo Zúñiga
Banco de México

Pedro N. Rodríguez
Sungard Trading and Risk Systems-BancWare

February 2008

Analysis of the Performance of Mexican Pension Funds: Evidence from a Stationary Bootstrap Application[*]

Arnulfo Rodríguez[†]
Banco de México

Gerardo Zúñiga[‡]
Banco de México

Pedro N. Rodríguez[§]
Sungard Trading and Risk Systems-BancWare

Abstract

This paper assesses the performance of Mexican pension funds (AFORES) by using an asset pricing model that includes macroeconomic factors and benchmark portfolios to explain returns. We apply a bootstrap statistical technique to obtain the cross-sectional distribution of performance measures (alphas) across all pension funds. This is done to determine whether a pension fund manager adds value to the portfolio before commissions charges, or if the performance observed, after controlling for the relevant factors, is simply explained by luck. Moreover, by comparing pension fund alphas to the distributions of alphas corresponding to lower rankings, we can find out if a particular fund statistically distinguishes itself from others. Our results provide evidence that pension funds managers do not add value to the portfolio and that funds are not distinguishable from each other.

Keywords: Pension funds, Performance evaluation, Stationary bootstrap.

JEL Classification: C14, G11, G23

Resumen

Este documento analiza el desempeño de las administradoras de fondos de ahorro para el retiro (AFORES) utilizando un modelo de fijación de precios que incluye factores macroeconómicos y portafolios de referencia para explicar los rendimientos. Utilizamos una técnica de remuestreo (bootstrap) para obtener una distribución de corte transversal de medidas de desempeño (alfas) para todas las AFORES en nuestra muestra. Esto se hace para determinar si un administrador de fondos de pensiones agrega valor a su portafolio (antes del cobro de comisiones) o bien, si después de controlar por otros factores relevantes, el desempeño es producto de la suerte. Al comparar la medida de desempeño (alfa) estimada para un fondo en particular con la distribución de alfas en rankings inferiores, podemos encontrar si un fondo se distingue de los demás. Nuestros resultados sugieren que los administradores de fondos de pensiones no agregan valor a su portafolio y que no pueden distinguirse unos de otros.

Palabras Clave: Fondos de pensiones, Evaluación de desempeño, Remuestreo estacionario.

1. Introduction

It has been ten years since the pension system for workers of the formal private sector in Mexico was reformed from a pay-as-you-go (defined benefit) scheme into a fully funded (defined contribution) system with individual accounts. These accounts consist of mandatory contributions by workers, employers and the government, and are administered and invested by specialized firms called AFORES (*Administradoras de Fondos de Ahorro para el Retiro*).

One characteristic of pension systems based on individual accounts is that workers' retirement income depends on the accumulated funds of their accounts. These funds grow over time based on several worker-specific variables such as age, income, contribution rate, and working life. The growth of these funds also depends on the rate of return obtained by pension fund managers and commissions charged for management and investment services. Therefore, it is of great relevance to assess the relative performance of AFORES regarding the returns on funds under management.

The main purpose of this study is twofold. First, we determine whether a pension fund manager adds value to the portfolio relative to a specific benchmark that represents a simple passive investment strategy, and if this performance can be distinguished from a random outcome (luck). Second, we use the bootstrap methodology to discriminate between AFORE-specific performance and the random outcomes associated with lower rankings to which other AFORES belong. This would allow us to differentiate one AFORE from the others based on its performance.

Our analysis consists of two parts. First we use an arbitrage pricing theory (APT) specification of returns to derive what is known in the literature as the Jensen's alpha.[1] In the second stage, we apply a stationary bootstrap technique to obtain the cross-sectional distribution of performance measures across all pension funds. We proceed to compare the alphas estimated in the first stage and the distribution obtained in the second stage to distinguish the estimated

---

[1] Jensen's alpha measures the excess return of a stock or portfolio above the return predicted by a pricing model given the stock or portfolio's exposure to risk (measured by the *betas* of the model) and the average market return. If this value is positive, then we say that the fund has earned excess returns which reflect superior investment abilities.

performance measure from random outcomes (luck) and to evaluate if AFORES can differentiate from each other based on these performance measures.

Previous literature on the performance assessment of Mexican pension funds is found in García-Verdú (2005). The author proposes that a two-factor model adequately captures the returns on Mexican pension funds. However, neither of those two factors is a fundamental economic variable.[2] His analysis suggests that no pension fund manager has superior asset-picking skills given a benchmark of government bonds.

Kosowski *et al.* (2006) conduct the first comprehensive examination of U.S. mutual funds performance that explicitly controls for luck and use a bootstrap statistical technique to generate the joint distribution of performance measures across all funds. Their main results indicate that a sizable minority of managers show superior asset-picking skills that add value to the portfolios under management and that this performance can be distinguished from luck.

To the authors' knowledge, not only is this work the first performance assessment of Mexican AFORES portfolios that explicitly controls for luck, but also the first that uses fundamental economic variables to explain gross returns on these funds. Our results suggest that, in the 2001-2007 period, AFORES do not add value to the funds under management relative to a set of benchmarks that represent a simple passive investment strategy, i.e. estimated alphas are not significantly different from zero.[3] The results of this first stage neither allow us to distinguish performance from random outcomes (luck) nor to make it possible to compare AFORES to each other in terms of their ability to add value to their portfolio. The results from the second stage confirm that the estimated alphas cannot be distinguished from random outcomes and also suggest that there are no significant differences among AFORES in terms of adding value to portfolios.

---

[2] Elton *et al.* (1995) demonstrate the importance of the inclusion of fundamental economic variables to explain bond returns. In our work this is relevant because AFORES were only allowed to invest in Mexican fixed-income assets before the year 2005. Moreover, García-Verdú (2005) mentions that performance measures are very similar when using the CAPM and a two-factor model. He suggests that this could be a result of not having a model that contains all the relevant factors in explaining pension fund returns.

[3] We also perform the analysis with a subsample for the period January 2005-July 2007 to focus on the period when investment in stocks was allowed. Our results remain qualitatively the same.

These results highlight the importance of providing incentives to pension fund managers to distinguish themselves from the rest by actively seeking higher returns. Ultimately, the provision of the right incentives would result in an improvement in workers' retirement accounts.[4]

In the next section we provide an overview of the current pension fund system. Section 3 introduces the APT model and describes the data. Section 4 presents the results of the estimation of the APT model. Section 5 introduces the stationary bootstrap technique along with the performance assessment statistically inferred from the cross-sectional distributions of performance measures (alphas). Finally, Section 6 presents our concluding remarks.

2. The Pension System of Individual Accounts in Mexico

The 1997 reform of the pension system in Mexico introduced individual accounts for every worker in the formal private sector. The Social Security Law enacted in 1997 established compulsory contributions to a retirement account from workers, employers and the federal government amounting to 6.5% of every worker base salary. In addition, the federal government contributes with 5.5% of the minimum wage to every worker's account, regardless of their level of income.[5]

Individual accounts are administered by AFORES and invested by specialized investment subsidiary entities called SIEFORES (*Sociedades de Inversión Especializadas en Fondos de Ahorro para el Retiro*). When a worker chooses (or is assigned to) a particular AFORE, the funds in the individual account are invested by that AFORE's SIEFORE.

SIEFORES are subject to an investment regime determined by the regulatory authority, CONSAR (*Comisión Nacional del Sistema de Ahorro para el Retiro*), which among other things sets limits to the risk that each SIEFORE can undertake with its investment choices. The

---

[4] Measuring the performance of pension funds in terms of returns net of commission charges is, perhaps, more appropriate than using gross returns since net returns reflect more accurately the value added by AFORES to workers' retirement accounts. To the extent that pension fund managers maximize this measure, it will result in higher pensions for workers when they retire. Ideally, we would follow the same two-stage procedure discussed above to measure investment abilities using net returns. However, there is no data on the selected benchmarks that take into account management fees and other commissions. Therefore, comparing AFORES' net returns with the benchmarks' gross returns might not be appropriate since we could be underestimating the value of the alphas and the interpretation of the results would be problematic.

[5] This subsidy represents a larger proportion of a worker's salary as the salary decreases. Furthermore, the contribution is contingent on the worker's continued participation in the system and is not subject to commission charges.

investment regime has been gradually modified to allow greater diversification while maintaining risk levels within certain limits.

Between 2002 and 2004 important changes were introduced to the investment regime. These changes were related to the following:

1) Since 1997 a minimum of 65% of instruments had to be government bonds with a maximum maturity (or revision) period of 183 days.[6] This restriction was first modified through an increase in the average maturity period to 900 days (December 2001).[7] Subsequently, the regulation was based on a daily valuation of risk (a VaR measure) in effect since November 2002.[8]

2) Originally, the type of issuer was restricted to the following limits:[9]

    a. Federal Government (minimum 65%)

    b. Private Issuers (maximum 35%)

    c. Financial Intermediaries (maximum 10%)

These limits were removed and replaced by creditworthiness (with a minimum rating of A) as measured by credit rating agencies. In addition, instruments issued by local and state governments and government-owned companies were allowed. [10]

3) The use of derivatives was introduced (November 2002). [11]

4) Investment in stocks was allowed up to 15% of total assets, as long as they tracked stock indices (May 2004).[12]

5) Exclusion of foreign issuers was lifted subject to a maximum share of 20% (this share includes both bonds and stocks) as long as the issuers were under the regulation of the

---

[6] *Circular CONSAR 15-1, Diario Oficial de la Federación de Junio 30 de 1997.*

[7] *Circular CONSAR 15-5, Diario Oficial de la Federación de Diciembre 5 de 2001.*

[8] The VaR limit was set at 0.6% over total net assets. *Circular CONSAR 15-8, Diario Oficial de la Federación de Noviembre 29 de 2002.*

[9] *Circular CONSAR 15-1*

[10] *Circular CONSAR 15-6, Diario Oficial de la Federación de Abril 8 de 2002.*

[11] *Circular CONSAR 15-8*

[12] *Circular CONSAR 15-12, Diario Oficial de la Federación de Mayo 26 de 2004.*

technical committee of the International Organization of Securities Commissions (IOSCO) and/or the European Union (May 2004).[13]

6) In May 2004 two investment funds were created based on life-cycle considerations:

    a. SIEFORE *Básica* 1 for workers with at least 56 years of age. This fund can only be invested in bonds (domestic and foreign (up to 20%)) and at least 51% is required to have some sort of inflationary protection.[14]

    b. SIEFORE *Básica* 2 for workers under 56 years of age. This fund allows investment in stocks (up to 15%) and up to 20% in foreign instruments (bonds or stocks).[15]

In 2007 a new set of changes were introduced to the investment regime, although the changes will not take effect until March 2008 and, thus, do not affect our analysis. These changes included:

-   Three new investment funds will be allowed (SIEFORES 3, 4 and 5) with investment limits based on life-cycle considerations.

-   SIEFORES 2, 3, 4 and 5 will be allowed to invest in *Notas*[16], variable rate instruments, and Real Estate and Infrastructure Trusts (*Fideicomisos de Infraestructura y Bienes Raíces*, FIBRAS).

The previously discussed modifications were gradually implemented and they occurred throughout the period under analysis in this paper. Since pension funds initially could only invest in a restricted subset of fixed-income assets as described above, our model utilizes two benchmarks that satisfy the investment regime restrictions throughout the period of analysis.

---

[13] *Ibid.*

[14] These instruments can either be denominated in "*Unidades de Inversión*, UDIs" (inflation-indexed units) or must guarantee a return equal or greater than the variation in UDIs.

[15] The VaR limit for SIEFORE *Básica* 2 was set at 1% of total net assets.

[16] Domestic and foreign debt instruments with the principal protected until maturity and linked to one or more of the permitted equity indices.

3. Model and Data

3.1 The Asset Pricing Model

In modern finance theory there is a fundamental interest in studying the relationship between returns and risks of assets. One of the most widely used models to analyze this relationship is the Arbitrage Pricing Theory (APT, Ross (1976)). The model incorporates different sources of non-diversifiable risk in a linear fashion and provides intuitive equilibrium conditions where no arbitrage is possible.

Following Burmeister *et al.* (2003) we can characterize the APT model as the combination of two basic equations:

$$r_{it} - E[r_{it}] = \beta_{i1} f_{1t} + ... + \beta_{iK} f_{Kt} + \varepsilon_{it} \tag{1}$$

where

r_{it} is the return on asset i at the end of period t;

$E[r_{it}]$ is the expected return on asset i;

β_{ij} is the risk exposure of asset i to the risk factor j, for $j = 1, …, K$;

f_{jt} is the value of the realization of the j-th risk factor at the end of period t, for $j = 1,…, K$;

ε_{it} is the end-of-period asset specific shock.

Equation (1) basically postulates that the difference between the actual return and the expected return for any asset (or portfolio) is equal to the sum of the asset specific shock and the exposures (betas) of that asset to the K risk factors multiplied by the realization of the K risk factors.

The APT model assumes there are no arbitrage opportunities for investors, i.e. an investor cannot construct an investment portfolio that will yield a sure profit without taking on risk or investing additional resources. It can be shown that if this condition is satisfied, then the second basic equation can be written as:

$$E[r_{it}] = \lambda_0 + \beta_{i1} \lambda_1 + ... + \beta_{iK} \lambda_K \tag{2}$$

6

Equation (2) postulates that there exist $K+1$ prices of risk, $\lambda_0, \lambda_1,…, \lambda_K$, at least one of them not zero, such that the expected return of asset i is equal to the sum of the prices of risk multiplied by the exposure of the asset i to the risk factor j (for $j = 1, …, K$).

Finally, substituting equation (2) in equation (1) we obtain the APT equation:

$$r_{it} - \lambda_0 = \beta_{i1}[\lambda_1 + f_{1t}] + ... + \beta_{iK}[\lambda_K + f_{Kt}] + \varepsilon_{it} \tag{3}$$

This equation can be empirically estimated using different methodologies. Burmeister *et al.* (2003) consider that there are three approaches available to empirically estimate the APT model:

1. The unobservable risk factors $f_1, …, f_K$ can be estimated using the statistical technique of principal components.

2. Some well-diversified portfolios may be constructed to approximate the K risk factors.

3. Economic theory and empirical evidence can be used to approximate the K risk factors using observable economic variables.

The first method is useful to find the number of relevant factors that explain most of the variation in returns. However, the resulting factors might not have an intuitive economic interpretation. Moreover, in this method there is no room for a constant, which in our case is the parameter of interest.

The second method is useful when there are clear investment strategies available to the investor that span the feasible set of investment alternatives (for example, small capitalization and large capitalization portfolios). However, finding the right combination of assets that span the feasible set of investments and approximate the risk factors is difficult.

The third method uses economic theory and empirical evidence to identify economic variables with intuitive interpretations. Moreover, this method uses additional economic information to explain returns, whereas methods 1 and 2 use returns data to explain returns. Since method 1 does not include a constant (our parameter of interest) and given the investment regime for SIEFORES (which limits the application for method 2), we consider that using the third approach is more appropriate for our objectives.

Following Chen *et al.* (1986), we assume that returns are generated by a mixture of benchmark portfolios and fundamental economic variables. We adapt the return generating process in Elton *et al.* (1995) and write it as

$$r_{it} = E[r_i] + \sum_{j=1}^{J} \beta_{ij}(R_{jt} - E[R_j]) + \sum_{k=1}^{K} \gamma_{ik} f_{kt} + \eta_{it} \qquad (4)$$

where

r_{it} is the return on pension fund i at time t;

R_{jt} is the return on benchmark portfolio j at time t;

f_{kt} is the unexpected change in the k-th fundamental economic variable at t;

β_{ij} is the sensitivity of pension fund i to the innovation of the j-th benchmark portfolio;

γ_{ik} is the sensitivity of pension fund i to the innovation of the k-th fundamental economic variable;

η_{it} is the time-t return of pension fund i that is unrelated to either benchmark portfolios or fundamental economic variables;

$E[\cdot]$ denotes expectation;

$E[f_k] = E[\eta_i] = 0$.

Note that f_k represents unexpected changes in a fundamental economic variable. By definition, the expected value of an unexpected change is zero or $E[f_k] = 0$.

From the Arbitrage Pricing Theory (APT) of Ross (1976), equation (4) leads to the following expression for the expected return on pension fund i:

$$E[r_i] = \lambda_0 + \sum_{j=1}^{J} \beta_{ij} \lambda_j^* + \sum_{k=1}^{K} \gamma_{ik} \lambda_k \qquad (5)$$

where

λ_0 is the return on the risk-free asset (R_F);

λ_j^* is the market price of sensitivity to the *j*-th benchmark portfolio;

λ_k is the market price of sensitivity to the *k*-th fundamental economic variable.

When variables in the return-generating process are benchmark portfolios, the APT market price of risk associated with such portfolio is the portfolio's expected return minus λ_0. Thus $\lambda_j^* = E[R_j] - \lambda_0$ for $j = 1,...,J$. In our case $J = 2$.[17] Substituting this expression into equation (5) and recognizing that $\lambda_0 = R_F$ yields:

$$E[r_i] = R_F + \sum_{j=1}^{J} \beta_{ij}\left(E[R_j] - R_F\right) + \sum_{k=1}^{K} \gamma_{ik}\lambda_k \qquad (6)$$

Following Elton *et al.* (1995), we substitute equation (6) into equation (4) and allow R_F to vary over time to obtain:

$$r_{it} - R_{Ft} = \sum_{j=1}^{J} \beta_{ij}\left(R_{jt} - R_{Ft}\right) + \sum_{k=1}^{K} \gamma_{ik}\left(\lambda_k + f_{kt}\right) + \eta_{it} \qquad (7)$$

Rearranging equation (7) yields:

$$r_{it} - R_{Ft} = \alpha_i + \sum_{j=1}^{J} \beta_{ij}\left(R_{jt} - R_{Ft}\right) + \sum_{k=1}^{K} \gamma_{ik} f_{kt} + \eta_{it} \qquad (8)$$

where

$$\alpha_i = \sum_{k=1}^{K} \gamma_{ik}\lambda_k \qquad (9)$$

We estimated Equation (8) for every pension fund to explain both gross nominal excess returns and net nominal excess returns (after considering commission charges).

---

[17] García-Verdú (2005) uses a two-factor model which consists of two benchmark portfolios, PiP-Guber and PiPG-Real, and finds that performance measures are quite similar between the one-factor (with the Pip-Guber as the only factor) and the two-factor models. However, it must be noted that PiP-Guber includes other assets that were not available to the SIEFORES given the restrictions of the investment regime and, therefore, is not an appropriate benchmark.

3.2 Data

The sample period goes from January 2001 to July 2007. This period was chosen mainly due to the availability of consistent data on the benchmarks used (available only since January 2001).[18] Since 1997, the number of AFORES has changed due to exits, mergers and new entries. In our analysis, we only consider AFORES that have survived in the industry ever since the inception of the new pension system. This selection raises the possibility of survival bias since we are leaving out of our sample those AFORES that existed at some point but left afterwards.

Mitigating the survival bias problem would require considering four additional AFORES in our analysis.[19] However, technical difficulties arise if we include these short-lived institutions and the bootstrap methodology is applied. In Sections 4 and 5 we discuss these issues in more detail. Furthermore, we did not take into account ten recently created pension funds due to their small number of observations.[20] As of July 2007 our sample of AFORES represented 82% of total assets under management.

Our analysis involves the estimation of excess gross returns of the SIEFORES portfolios. We chose an index of government securities with a fixed rate and up to a 30-day maturity (CETES and BONDES) as the risk-free asset and the 12-month SIEFORES gross return as the AFORES portfolio returns. After January 2005 we consider returns of SIEFORE *Básica* 2.

In the estimation of excess returns we use an asset pricing model that introduces fundamental macroeconomic variables. The choice of these variables is based on both the characteristics of the Mexican economy and findings of a vast literature on the relationship between financial markets and macroeconomic variables (see Chen *et al.* (1986), Elton *et al.* (1995) and Burmeister *et al.* (2003)).

The variables used in our estimations are: 1) monthly nominal excess gross returns obtained from the historical 12-month nominal return of the SIEFORES published by CONSAR;

---

[18] We must note that since January 2001 only another AFORE was present in the industry (HSBC) but unfortunately for the years 2002 and 2003 there were no available data.

[19] Among those four, AFORE Allianz Dresdner was the last one to exit the market (December 2003). This implies having left thirty six observations out of our sample for this particular AFORE.

[20] ING was the AFORE of recent creation with the largest number of observations (sixty nine) that was left out of the analysis.

2) nominal returns of an index of government securities with a maximum of 30-day maturity and fixed rate (PiPG-Fix1M) published by an authorized Mexican price vendor and valuation firm (*Proveedora Integral de Precios*, PiP); 3) two benchmark indices published by PiP: BONDES 182 (PiPG-Bonde182) and another that consists of UDIBONOS (PiPG-Udibonos); 4) a measure of investors risk appetite provided by Credit Suisse First Boston; 5) a measure of time horizon risk derived from the differences in the returns on 5-year government bonds and 30-day government bonds, these were calculated from indices published by PiP (PiPG-Fix5A and PiPG-Fix1M respectively); 6) a measure of inflation risk derived from inflation expectations for the next 12 months which were surveyed by Banco de Mexico[21]; 7) a measure of real sector activity derived from GDP expectations for the following year which was also surveyed by Banco de Mexico; 8) monthly returns on the IPC (*Indice de Precios y Cotizaciones*) published by the Mexican Stock Exchange (*Bolsa Mexicana de Valores*); 9) changes in the oil price of the Mexican export blend (*Mezcla Mexicana de Exportación*), available from Banco de Mexico; and 10) changes in the yield of the 10-year US Treasury Bond published by the Federal Reserve Board.

Table I: Dependent Variables

SERIES	VARIABLE	SOURCE
Gross Returns 02/2001-10/2006	The monthly change in an index constructed with return data (12-months) published by CONSAR for the following AFORES: Banamex, Bancomer, Banorte, Inbursa, Principal, Profuturo, Santander, and XXI.	CONSAR.

Table II: Benchmarks

SERIES	VARIABLE	SOURCE
Benchmarks UDIBONOS 02/2001-10/2006 (coefficient: β_{i1})	Monthly changes in the PiPG-Udibonos index.	Proveedora Integral de Precios (PiP).
Benchmarks BONDES 182 02/2001-10/2006 (coefficient: β_{i2})	Monthly changes in the PiPGBonde-182 index.	Proveedora Integral de Precios (PiP).

---

[21] The inflation expectations data were obtained from the 2001-2007 monthly surveys entitled "*Encuesta sobre las expectativas de los especialistas en economía del sector privado*"
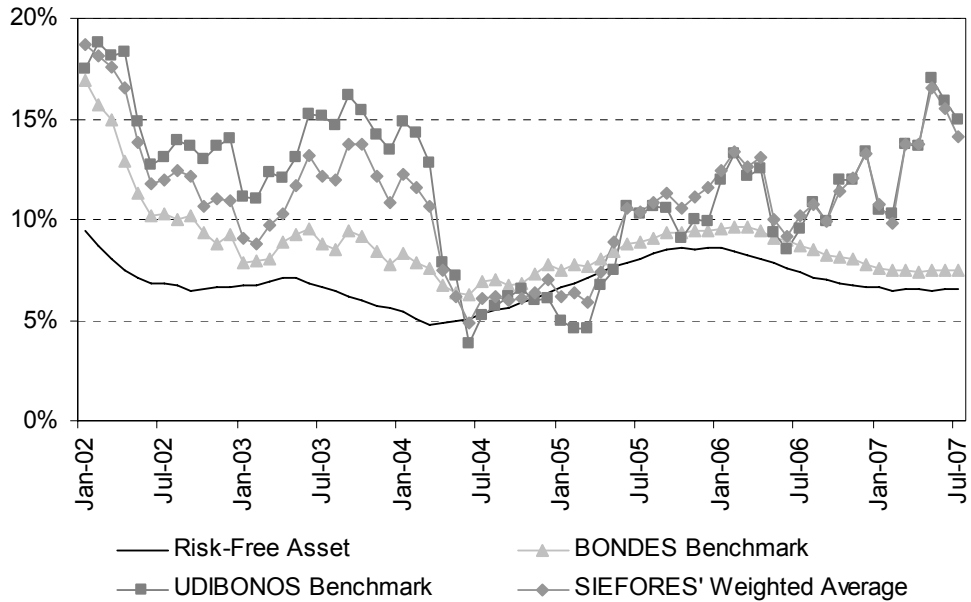
Table III: Macroeconomic Variables

SERIES	VARIABLE	SOURCE
Economic Cycle Risk 02/2001-10/2006 (coefficient: γ_{i1})	This is measured as the monthly change in the forecast of GDP for the following year.	Survey of Expectations of the Private Sector. Banco de México.
Inflation Risk 02/2001-10/2006 (coefficient: γ_{i2})	This is measured as the forecast error: the difference between forecasted and observed inflation.	Survey of Expectations of the Private Sector. Banco de México.
Confidence Risk 02/2001-10/2006 (coefficient: γ_{i3})	Changes in the Global Risk Appetite Index published by Credit Suisse First Boston. This index aims to capture investors' confidence, by measuring changes in the relative performance of safe assets (government bonds) versus volatile assets like equities and emerging country bonds.	Credit Suisse First Boston.
Market Risk 02/2001-10/2006 (coefficient: γ_{i4})	This is approximated with the monthly changes in the Mexican Stock Exchange Index.	Banco de México.
Time Horizon Risk 02/2001-10/2006 (coefficient: γ_{i5})	This variable is obtained by subtracting short-term (30-days) government bond yields from long-term (5 years) fixed rate government bonds yields.	PiP.
Oil 02/2001-10/2006 (coefficient: γ_{i6})	Monthly changes in the price (dollars per barrel) of the Mexican Oil Export Mix.	Banco de México.
Foreign Interest Rate 02/2001-10/2006 (coefficient: γ_{i7})	Monthly changes in the yield of 10-year US Treasury Bonds.	US Federal Reserve.

Figure I: Portfolio Composition

- Public Debt Instruments
- Private Debt Instruments
- Domestic Equities
- Foreign Debt Instruments
- Foreign Equities
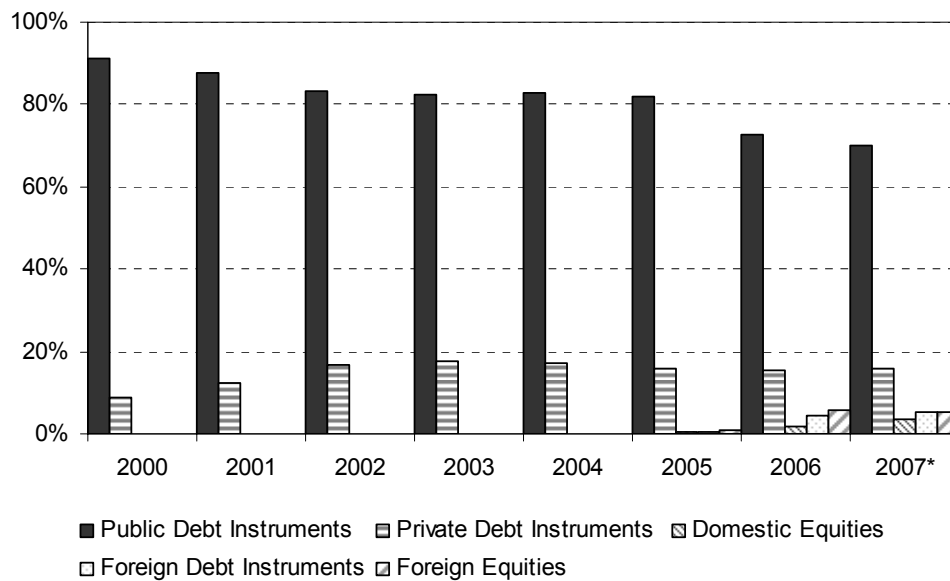
* July 2007.
Source: CONSAR.

Figure II: Gross Returns
(12-months)

Source: PiP and CONSAR.

One clarification note regarding the benchmark indices is in order. Given the high concentration of government bonds in the composition of pension funds portfolios, we used two benchmark indices that represent a simple passive investment strategy that included only government bonds. Both BONDES 182 and UDIBONOS were investment vehicles available to pension fund managers subject to the past investment regime. Between January 2001 and the creation of SIEFORES 1 and 2 in May 2004, pension fund managers could have invested up to 100% of their assets in government bonds with at least 51% invested in instruments with inflationary protection (either denominated in inflation-indexed units, UDIs, or with a guaranteed return equal or greater than the variation in the value of UDIs). This restriction is satisfied by UDIBONOS and BONDES 182.

There were also restrictions on the maturity period of instruments: the limit was 183 days from July 1997 to December 2001. Between the latter and November 2002 the limit increased to an average of 900 days. From November 2002 onwards the restriction was shifted to a VaR measure.[22] These restrictions are also satisfied by BONDES 182 and UDIBONOS. After the

---

[22] *Circulares CONSAR 15-1, 15-5, 15-8.*

13

separation of portfolios in May 2004 into SIEFORES 1 and 2 we only analyze the returns of SIEFORE *Básica* 2. This fund could be invested up to 100% in government bonds with no restrictions regarding inflation-indexed instruments or maturity periods.

4. Results

Table IV shows the estimation results obtained with OLS, whose parameters' standard errors are adjusted for potential autocorrelation and heteroskedasticity by the method proposed by Newey and West (1987).[23] As can be seen from Table IV below, none of the alphas is significantly different from zero. In other words, according to this evidence the managers of such AFORES do not add value to their portfolio for the sample analyzed.

The mitigation of the survival bias problem requires including those AFORES that were present for a very short period of time between 2001 and 2007. However, if we had done their estimation we would have obtained relatively high standard errors for their alpha due to the low number of observations. Kosowski *et al.* (2006) mention that a fund having a short life will have a high variance estimated alpha distribution and alphas for such funds will tend to be spurious outliers in the cross-section.

---

[23] Ferson *et al.* (2003) use the Newey-West (1987) method when they have autocorrelation in the explanatory variables and/or heteroskedasticity in the estimated regression.

**Table IV: Estimated coefficients obtained with OLS regressions for every
pension fund excess gross returns**

The following regression model is estimated on excess gross returns using monthly data from February 2001
to July 2007:

$$r_{it} - R_{Ft} = \alpha_i + \sum_{j=1}^{J} \beta_{ij}\left(R_{jt} - R_{Ft}\right) + \sum_{k=1}^{K} \gamma_{ik} f_{kt} + \eta_{it}$$

The p-values are below the coefficients and were corrected for heteroskedasticity and autocorrelation using
the Newey-West estimator. R^2s from each regression are reported in decimal form.

Nominal Excess Gross Returns
Benchmark: UDIBONOS and BONDES 182 Combination

AFORE	α_i	γ_{i1}	γ_{i2}	γ_{i3}	γ_{i4}	γ_{i5}	γ_{i6}	γ_{i7}	β_{i1}		β_{i2}		R^2
Bancomer	-0.07	-0.38	-0.41	-0.07	0.01	0.08	0.01	-0.52	0.31	***	0.58	**	0.45
p-value	0.68	0.20	0.25	0.37	0.61	0.50	0.62	0.20	0.00		0.02		
Profuturo	-0.08	-0.22	-0.42	-0.01	0.02	0.07	0.02	-1.14 **	0.44	***	0.35		0.51
p-value	0.68	0.48	0.28	0.88	0.53	0.59	0.50	0.02	0.00		0.16		
XXI	-0.10	-0.38	-0.44	-0.08	0.02	0.12	0.02	-0.49	0.34	***	0.46	*	0.48
p-value	0.64	0.26	0.26	0.42	0.49	0.34	0.40	0.25	0.00		0.09		
Banamex	-0.10	-0.29	-0.33	-0.06	0.02	0.11	0.02	-0.76	0.49	***	0.33		0.58
p-value	0.56	0.30	0.39	0.49	0.48	0.37	0.34	0.13	0.00		0.15		
Principal	-0.13	-0.26	-0.51	-0.12	0.02	0.12	0.02	-0.46	0.37	***	0.42	*	0.50
p-value	0.51	0.40	0.16	0.20	0.54	0.26	0.43	0.24	0.00		0.09		
Banorte	-0.19	-0.25	-0.36	-0.04	0.01	0.12	0.02	-0.53	0.28	***	0.43	*	0.46
p-value	0.27	0.36	0.29	0.63	0.76	0.23	0.34	0.13	0.00		0.06		
Santander	-0.20	-0.34	-0.57	-0.04	0.02	0.13	0.02	-0.69	0.35	***	0.33		0.50
p-value	0.34	0.26	0.16	0.63	0.52	0.24	0.46	0.12	0.00		0.21		
Inbursa	-0.28	-0.49 *	-0.36	-0.16 **	0.00	0.00	0.01	-0.04	0.08		0.58	*	0.24
p-value	0.27	0.06	0.32	0.04	0.88	0.99	0.54	0.92	0.50		0.06		

*, ** and *** denote significance of 10, 5 and 1 percent, respectively.

We also estimated equation (7) to explain both the cross-sectional variation and time series
of gross returns. This estimation was done by using the iterated non-linear seemingly unrelated
regressions (ITNLSUR) discussed by Gallant (1987). Since equation (7) is equivalent to equation
(8) with restriction (9) imposed, we compare the results of estimating equation (8) alone vs.
estimating equation (7) to determine whether imposing the APT restriction (9) on our nine-factor
model results in a statistically reduction in explanatory power. We find that the APT restriction
holds. Table V shows the results.

Table V: Likelihood ratio test statistics to evaluate the APT model restriction

We compare the APT model against the non-restricted model. The dependent sample consists of 8 AFORES. The sample periods are from February 2001 through July 2007 and from January 2005 through July 2007. Null hypothesis: the APT model restriction (equation (9)) holds; alternative hypothesis: the APT restriction does not hold. Small-sample adjustment (see Gallant (1987)): Reject null when $L > qF_\alpha$, where F_α is F statistic at α level of significance with q degrees of freedom in numerator and $nM - p$ degrees of freedom in denominator. $n = 77$ and $n = 31$ observations for the complete sample and subsample, respectively. $M = 8$ equations, and α = 0.05.

				Pairwise Test		
Benchmark	L^a	p^b	q^c	$qF_{0.05}$	L - qF0.05	Reject the null that the APT holds?
Complete sample	0.12	80	8	15.52	-15.40	No
Subsample	0.00	80	8	15.52	-15.52	No

[a] $L = n(\ln|\Sigma_1| - \ln|\Sigma_2|)$, where Σ_1 is the variance-covariance matrix of restricted residuals and Σ_2 is the variance-covariance matrix of unrestricted residuals.
[b] p = total number of estimated parameters.
[c] q = number of restrictions.
  n = number of observations.

In an attempt to find out if the APT model with only the two benchmark portfolios has explanatory power equal to the nine-factor APT model, we compared their estimation results by using the likelihood ratio test. For both APT models, we estimated equation (7) using ITNLSUR to explain both the cross-sectional variation and time series of gross returns. We find that the null hypothesis of the APT model with only the two benchmark portfolios does not hold. Table VI shows the results.

Table VI: Likelihood ratio test statistics to evaluate the model consisting of only benchmark portfolios as risk factors

We compare the APT model with only the two benchmark portfolios to the nine-factor APT model. The dependent sample consists of 8 AFORES. The sample periods are from February 2001 through July 2007 and from January 2005 through July 2007. Null hypothesis: the APT model with only the two benchmark portfolios holds; alternative hypothesis: the APT model with only the two benchmark portfolios does not hold. Small-sample adjustment (see Gallant (1987)): Reject null when $L > qF_\alpha$, where F_α is F statistic at α level of significance with q degrees of freedom in numerator and $nM - p$ degrees of freedom in denominator. $n = 77$ and $n = 31$ observations for the complete sample and subsample, respectively. $M = 8$ equations, and α = 0.05.

				Pairwise Test		
Benchmark	L^a	p^b	q^c	$qF_{0.05}$	L - qF0.05	Reject the null that the restricted model holds?
Complete sample	129.26	79	63	82.53	46.73	Yes
Subsample	154.07	79	63	82.53	71.54	Yes

[a] $L = n(\ln|\Sigma_1| - \ln|\Sigma_2|)$, where Σ_1 is the variance-covariance matrix of restricted residuals and Σ_2 is the variance-covariance matrix of unrestricted residuals.
[b] p = total number of estimated parameters.
[c] q = number of restrictions.
  n = number of observations.

5. The Stationary Bootstrap Technique and Cross-Sectional Distributions of Performance Measures (alphas)

In this section we apply a bootstrap statistical technique to determine whether a pension fund is good at adding value to their customers portfolio before commissions charges, or whether it is simply lucky.

We use the stationary bootstrap method proposed by Politis and Romano (1994). This traditional resampling algorithm was used to generate cross-sectional distributions of alphas. In building such distributions, we impose the condition that the true alphas are zero for every pension fund with a given rank. Kosowski *et al.* (2006) impose such condition to determine whether a mutual fund is lucky (unlucky) or good (bad) at asset-picking. In constructing their artificial return series, they impose the null hypothesis of zero true performance ($\alpha_i = 0$).

Formally, in our model, when strictly following Kosowski *et al.* (2006) we would generate the following pseudo time series of excess returns (er_{it}^b) for each pension fund:

$$er_{it}^b = \sum_{j=1}^{J} \hat{\beta}_{ij}\left(R_{jt} - R_{Ft}\right) + \sum_{k=1}^{K} \hat{\gamma}_{ik} f_{kt} + \hat{\eta}_{it}^b \tag{10}$$

where $\hat{\eta}_{it}^b$ are the residuals resampled with replacement from the vector of residuals obtained in the estimation with the original sample. Since the stationary bootstrap resamples observations (not residuals) with replacement in blocks of random size, we must find an equivalent condition for imposing the null hypothesis of zero true performance.[24] Such equivalence is found when

$$er_{it}^b - \hat{\alpha}_i = \hat{\alpha}_i - \hat{\alpha}_i + \sum_{j=1}^{J} \hat{\beta}_{ij}\left(R_{jt} - R_{Ft}\right) + \sum_{k=1}^{K} \hat{\gamma}_{ik} f_{kt} + \hat{\eta}_{it}^b \tag{11}$$

In other words, subtracting the estimated alpha (obtained with the original sample) from resampled observations of excess returns is equivalent to imposing the null hypothesis of alpha equal to zero for the case of residual resampling.

It is worth mentioning that the resampled observations preserve the contemporaneous correlation across all pension funds – i.e. in each bootstrap iteration, the block of resampled

---

[24] We decided to use the stationary bootstrap instead of the simple bootstrap procedure of residual resampling in order to take into account potential autocorrelation of excess returns.

observations correspond to the same time periods across all pension funds. This is done to reflect the fact that pension funds hold very similar portfolios during most of our sample due to the strict restrictions imposed by the investment regime.[25] However, this contemporaneous correlation preservation makes it very difficult (if not impossible) to deal with the survival bias problem for a couple of reasons. First, if we considered short-lived AFORES, we would have to resample observations whose total number would be equal to the number of observations associated with the shortest-lived AFORE. Second, assuming that the econometric estimations with the resampled observations could be done, there would be very high standard errors for the alphas and the tails of the cross-sectional distributions would contain many outliers.[26]

We ranked the funds according to their alpha obtained from the estimations done with the original data. Figure III shows the cross-sectional distributions of alphas for each rank.[27] The vertical lines in those figures indicate the estimated alphas from the estimations with the original data. Figure III can be interpreted as follows: the vertical line represents the estimated performance (alpha) for each pension fund resulting from the estimation of equation (8). The densities represent the distribution of alphas for each rank which were obtained from our bootstrapping method. Since the bootstrap method consists of a random resampling (with replacement) of observations, this distribution can be thought of as purely randomly generated values of performance (alphas). In other words, a value of alpha will not be attributed to luck when the area (to the left of the vertical line) is greater than 0.90 (p-value < 0.10) of the corresponding cross-sectional distribution.
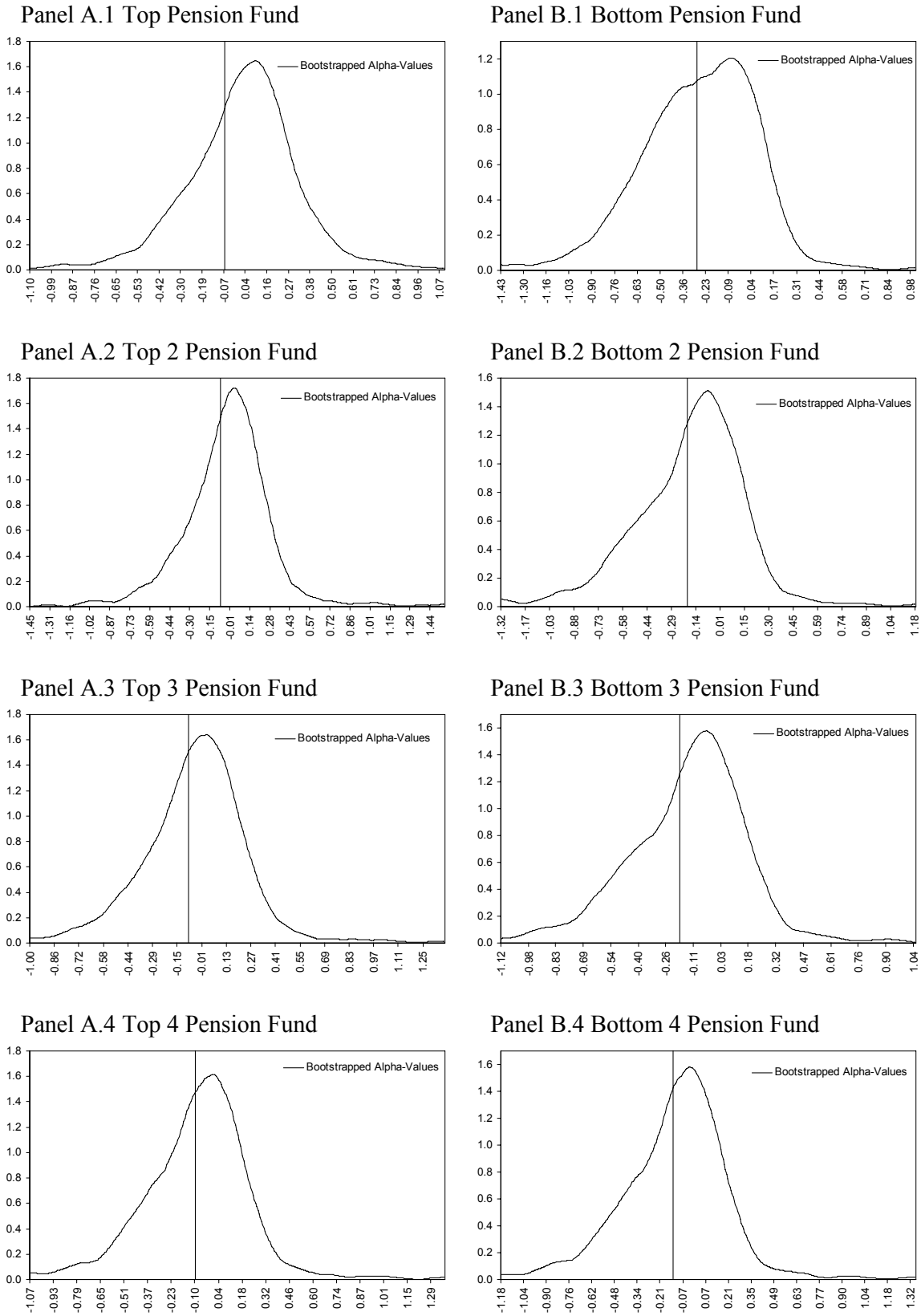
It can be seen from Figure III that our results provide strong evidence that no fund distinguishes itself from the distribution of random outcomes, and not even the highest-ranked fund has a significant superior performance (compared with the other funds) when gross returns are considered.

---

[25] Kosowski *et al.* also allow for the possibility of cross-sectional dependence among fund residuals that might, for example, be due to funds holding similar stocks at the same time.

[26] In fact, all pension funds would look like short-lived funds.

[27] For example, the top distribution consists of all the alphas that came out to be the highest from each bootstrap iteration.

Figure III: Cross-sectional distributions of alphas for nominal excess gross returns

Panel A.1 Top Pension Fund

Panel B.1 Bottom Pension Fund

Panel A.2 Top 2 Pension Fund

Panel B.2 Bottom 2 Pension Fund

Panel A.3 Top 3 Pension Fund

Panel B.3 Bottom 3 Pension Fund

Panel A.4 Top 4 Pension Fund

Panel B.4 Bottom 4 Pension Fund

We also compared pension fund alphas to the distributions of alphas corresponding to lower rankings. This exercise allows us to determine whether or not a particular pension fund is statistically distinguishable from the ranking assigned to other funds. If we determine no statistical difference, then both AFORES will be indistinguishable in terms of adding value to their portfolio. Table VIII shows the p-values from comparing pension fund alphas to the distributions of alphas corresponding to lower rankings. In this table, the first column contains the AFORES analyzed and which are ordered in a descending way according to its estimated alpha (from the original sample). The row with numbers 1, 2, 3, ….., 8 indicate the respective rankings of cross-sectional distributions with which the estimated alphas are compared. For example, the diagonal elements correspond to the cases where a particular AFORE performance is compared with the cross-sectional distribution corresponding to the same rank.

Table VII: P-values obtained from comparing pension fund gross return alphas to the distributions of alphas corresponding to equal or lower rankings

	1	2	3	4	5	6	7	8
Bancomer	0.682	0.615	0.556	0.519	0.48	0.44	0.393	0.308
Profuturo		0.627	0.565	0.533	0.486	0.453	0.403	0.312
Siglo XXI			0.594	0.556	0.517	0.476	0.431	0.332
Banamex				0.561	0.52	0.48	0.433	0.333
Principal					0.567	0.537	0.486	0.368
Banorte						0.615	0.573	0.44
Santander							0.595	0.449
Inbursa								0.542

For the case of gross returns, it can be seen from Table VII above that AFORES are not statistically distinguishable from one another according to the p-values obtained. The results for the subsample January 2005 through July 2007 do not qualitatively change the results obtained for the whole sample. However, had we included those AFORES that existed shortly between 2001 and 2007, it could have been possible to statistical distinguish at least some AFORES from some of those not included in our analysis.

6. Conclusions

In this paper we evaluated the performance of pension fund managers in Mexico, conditional on an asset pricing model with macroeconomic variables and two benchmark portfolios. Furthermore, we aimed at distinguishing the estimated values of performance (alphas) from random outcomes using a stationary bootstrap technique. In addition, we compared pension fund alphas to the distributions of alphas corresponding to lower rankings in order to distinguish between AFORES.

Our results for the January 2001 through July 2007 period suggest that pension fund managers do not seem to add value to their portfolio. Moreover, when comparing pension fund gross returns alphas to the distributions of alphas corresponding to lower rankings, we find that AFORES are not statistically distinguishable from one another.

These results should be interpreted with caution mainly because of three reasons. First, the estimated alphas are conditional on the specific APT model used and its inherent linearity assumption. Moreover, we acknowledge that the cross-sectional distributions generated by the bootstrap methodology might be different from those distributions obtained if all AFORES in the market (those who exited and those who entered in recent years) were included. Second, the recent changes in the investment regime may take more time to be reflected in pension fund returns. Third, one must take into account that pension fund managers face a long investment horizon. Indeed, workers that are members of the new generation with individual accounts will begin to retire in approximately 15 years.

It is worth mentioning that the regulatory authority (CONSAR) has recently modified the criteria for the allocation process of workers that do not choose a particular AFORE and for workers that want to switch to a different AFORE. Previously, workers were allowed to switch to a different AFORE only if they had been registered in one AFORE for at least one year or if the desired AFORE charged lower commissions. Similarly, CONSAR assigned workers who did not choose a particular AFORE to those administrators charging the lowest commissions. With the new regulation, both workers will be allowed to switch to a different AFORE and CONSAR will assign new workers based on an indicator of net returns. This modification strives to provide

AFORES with the incentives to maximize net returns and not only to focus on commissions. To the extent that the investment regime continues to evolve aiming to achieve an optimal combination of risk and returns and as long as pension fund managers take advantage of the deregulatory changes to this regime, it could be possible to observe an improvement in the performance of Mexican pension fund managers.

References

Burmeister, Edwin, Richard Roll, and Stephen A. Ross, 2003, Using Macroeconomic Factors to Control Portfolio Risk, mimeo, available at:

http://web.econ.unito.it/nicodano/Roll_Ross_APT_Portfolio_Management.pdf

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, 1986, Economic Forces and the Stock Market. *Journal of Business* 59, 383-403.

Davison, A. C., and D. V. Hinkley, 1997. *Bootstrap Methods and their Application* (Cambridge University Press).

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 1995, Fundamental Economic Variables, Expected Returns, and Bond Fund Performance. *Journal of Finance* 50, 1229-1256.

Elton, Edwin J., Martin J. Gruber, and Mustafa Gultekin, 1981, Expectations and Share Prices. *Management Science* 27, 975-987.

Ferson, Wayne E., Sergei Sarkissian and Timothy T. Simin, 2003, Spurious regressions in financial economics, *Journal of Finance* 58, 1393-1413.

Gallant, Ronald A., 1987. *Non-linear Statistical Models* (John Wiley & Sons, New York).

García-Verdú, Santiago, 2006, An assessment of the performance of the Mexican private pension funds, Mimeo, University of Chicago.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis. *Journal of Finance* 61, 2551-2595.

Kotlikoff, Laurence, 1999, The World Bank's Approach and the Right Approach to Pension Reform, Mimeo, Boston University.

Kotlikoff, Laurence, 2006, Pension Reform – The Triumph of Form over Substance? presented at the 2006 Latin American Meeting of the Econometric Society and the Annual Meeting of the Latin American and Caribbean Economic Association.

Newey, Whitney, and Kenneth West, 1987, A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55, 703-708.

Politis, Dimitris N., and Joseph P. Romano, 1994, The Stationary Bootstrap. *Journal of the American Statistical Association* 89, 1303-1313.

Ross, Stephen A., 1976, The arbitrage pricing theory of capital asset pricing, *Journal of Economic Theory* 13, 341-360.