Investigating Persistence in the US Mutual Fund Market: A Mobility Approach

KONSTANTINOS DRAKOS[*]

Athens University of Economics and Business

NICHOLAS GIANNAKOPOULOS

University of Patras

PANAGIOTIS TH. KONSTANTINOU[†]

Athens University of Economics and Business

Performance persistence in the US mutual fund market is investigated, modeling risk-adjusted performance as a Markov Chain. This allows us to explore whether there is a higher probability for funds to remain in their initial ranking, compared to the probability that funds exhibit some kind of movement. We find some degree of inertia due to non-uniformity of transition probabilities across states. Our analysis also assesses the proximity of empirical transition matrices to two benchmark matrices, identifying the no-persistence/perfect immobility cases. We find that the observed transition matrices are closer to the no-persistence benchmark and also that performance persistence has decreased over time.

Keywords: Markov Chain, Mobility, Mutual Funds, Persistence.

JEL: G14, G23

1 Introduction

Mutual funds' performance persistence, defined as an increased propensity for relative performance rankings to repeat between successive periods, has important implications for

---

market efficiency. In particular, market efficiency implies that risk-adjusted performance is serially uncorrelated and unpredictable (*i.e.* martingale difference). If these conditions were satisfied, risk-adjusted performance would exhibit no persistence, since past performance would contain no predictive power over future performance.

Performance persistence implies that the underlying structure of the fund market exhibits a certain degree of inertia, stemming from insufficient mobility (*i.e.* 'excess' tendency for repeat performers). In contrast, the EMH leaves no room for inertia, since risk-adjusted excess returns follow a martingale difference process that by default cannot be predicted (Fama 1965, 1970). In particular, under the null of no persistence in performance, we may define a notion of 'sufficient' mobility, boiling down to an equality of the proportion of funds staying in their past rankings to the proportion of funds moving across ranking positions. In other words, transitions to any direction (upwards, downwards or staying put) ought to represent equally likely outcomes.

Should we expect performance to persist and if yes at what horizon? Berk and Green (2004) argue that we may expect managers to possess a short-lived informational advantage. The short life of this advantage is an equilibrium outcome, when investors direct their capital towards recent winners. Our paper demonstrates that this is in fact the case during the 2000s.

Previous studies provide evidence that stock selection ability tends to be persistent over periods as short as one year. They find that although funds on average generate negative abnormal returns, *relative performance* persists (Grinblatt and Titman, 1992; Hendricks *et al.*, 1993; Goetzmann and Ibbotson, 1994; Brown and Goetzmann, 1995; Malkiel, 1995; Elton *et al.*, 1996; Gruber, 1996; Carhart, 1997; Bollen and Busse, 2004; Busse and Irvine, 2006). Most of these studies attribute persistence partially to managerial ability. On the other hand, Carhart (1997), argues that the superior performance of top funds is a result of the momentum anomaly (see Jegadeesh and Titman 1993). Including a momentum factor in his return model, Carhart finds that persistence is largely absent, except among the worst performers, where persistence is generated by persistently high expenses. This finding indicates that fund managers possess little stock selection skill, since the top performing funds generate their superior returns simply by holding stocks that have recently had high returns.

The present study revisits the issue of persistence in mutual fund performance, focusing on a measurement period of three years, allowing us to assess the existence and temporal evolution of medium-term performance persistence. We start by generating risk-adjusted excess performance for each three-year-period in our sample, based on the estimated abnormal return ('alpha') from Carhart's (1997) four-factor model that augments the Fama and French (1993) three-factor model by the Jegadeesh and Titman (1993) momentum

anomaly. We then rank funds by the estimated abnormal return and form quartiles. We then compute the abnormal return generated by each fund, and form again quartiles the following three years.[1] We provide evidence of statistically significant persistence in risk-adjusted excess returns (repeat performers), but we also document a temporal tendency for performance persistence to be reduced. That is we find that performance persistence has been declining over time.

We then propose to model successive performance rankings of US funds and their dynamic evolution as a Markov Chain, providing the generic mechanism from which the widely employed 'winner-loser' approach stems as a specific case. In particular, the Markovian structure we suggest, can be viewed as a generalization of the nonparametric, 2×2 contingency tables used in previous studies (Brown and Goetzmann, 1995; Brown *et al.*, 1999; Goetzmann and Ibbotson, 1994), which allows to characterize more *states of nature* than the simple 'winner-loser' dichotomy, in a manner analogous to the portfolio ranking techniques also employed in the literature. Our suggestion amounts to estimating the probabilities that a fund ranked in a particular class, will remain in the same class or will transit to a different class during the evaluation period. In this sense, our proposed method provides evidence on whether funds in a particular class during the ranking period improve or worsen their position and by what probability.[2] We should also highlight that our suggestion is a generalization of the 2×2 contingency tables approach to the $n \times n$ case.[3]

In a next step, we use the estimated transition probabilities to assess uniformity in a formal statistical manner. In our dataset, we find that uniformity is broadly rejected, while irreducibility holds.[4] In addition, we propose to evaluate the degree of performance persistence by means of a set of mobility indices, which have been previously used in the context of credit ratings (Bangia *et al.,* 2002; Jafry and Schuermann, 2004; Trueck, 2004; Trueck and Rachev, 2006) as well as the evolution of income distribution (Geweke *et al.,* 1986*;* Shorrocks, 1978). These indices allow us to assess the degree of inertia in the employed performance measures, hence providing evidence for the evolution of performance persistence over time. This is a key contribution of our work. Essentially we move away from

---

[1] This should not be taken to imply that in order for a fund to be included in the analysis, the fund must have survived for 72 (36+36) consecutive months. See the discussion in subsection 3.2.

[2] In the context of hedge funds performance, Brown et al. (1999) and Agarwal and Naik (2000) implicitly refer to this issue, which they do not pursue any further.

[3] As correctly pointed out by a reviewer, our approach amounts to a sub-class of the (general) contingency tables methodology, when the the contingency tables are square $n \times n$, relative to the general $n \times m$ case.

[4] Uniformity holds when states have equal probabilities of being visited regardlsess of the initial state. It also implies that there is no persistence in the relevant performance measures. Irreducibility implies that all states have a positive probability of being visited.

a simple evaluation of whether performance persistence is present in the data, and we show how performance persistence has evolved over time. The evolution of performance persistence across time-periods has been broadly overlooked in previous studies. Our findings indicate that the degree of performance persistence has declined over time, becoming much less pronounced, especially during the 2000s.

In order to evaluate the temporal evolution of performance persistence, we define two benchmarks transition matrices: the identity matrix, which represents perfect immobility of relative performance rankings (*complete persistence*); and the *no persistence matrix* (whose elements are equal to each other, i.e. uniformity holds). We then explicitly assess the level of (in-) efficiency by exploring the 'distance' of the empirical transition probabilities from the two benchmark theoretical transition probabilities, which provides valuable information about the degree of mobility (lack of persistence) and its temporal evolution thereof.

The rest of the paper is organized as follows. Section 2 discusses the model we employ to measure performance. Then once a performance metric is adopted, we explain how performance persistence can be cast in a Markovian framework, discussing also the mobility indices we employ. Section 3 discusses the utilized dataset and presents some first results regarding performance persistence, while section 4 contains our main empirical results. The last section concludes.

2 Persistence in the Mutual Fund Market: Implications for Efficiency

This section reviews our methodology for measuring performance persistence, and explains our approach for assessing performance persistence.

2.1 Mutual Fund Performance

The starting point in assessing performance persistence is to devise an appropriate measure of performance that adequately takes into account systematic risk. Previous studies have examined either stock selection or market timing ability. We focus on the former in this paper. Studies of stock selection, in the spirit of the early contribution by Jensen (1969), employ the intercept ('alpha') of factor model regressions as measures of abnormal returns, which follow from picking a portfolio of stocks that outperforms a risk-adjusted benchmark. Following this strategy, we assume that the return generating mechanism is given by Carhart's (1997) four-factor model which is an augmented version of the Fama and French (1993) three-factor model including the Jegadeesh and Titman (1993) momentum anomaly:

$$R_{ft} = \alpha_{ft} + \beta_{1f}RMRF_t + \beta_{2f}SMB_t + \beta_{3f}HML_t + \beta_{4f}MOM_t + u_{ft} \qquad (1)$$

where R_{ft} is the excess return of mutual fund f, $RMRF_t$ is the excess return on a value weighted aggregate market proxy, while SMB_t, HML_t, and MOM_t are the returns on value weighted zero-investment factor mimicking portfolios for size, book-to-market equity and one-year momentum in stock returns.

Another strand of the literature examines whether stock selection ability is persistent, basing tests of persistence on correlation of the period-to-period fund performance. In this paper, we rank funds on a three-year period based on risk-adjusted performance measures, following Carhart (1997), but we estimate the post-ranking performance also over a three year interval. In this sense we are able to evaluate the degree that ability persists over longer horizons. In addition, using non-overlapping three-year intervals, we are also able to assess the evolution, if any, of ability persistence.

2.2 Persistence and Mobility: A Markov Chain Framework

In this subsection we explain how it is possible to cast the dynamics of performance persistence in a more general Markovian setting. Consider F mutual funds with a typical fund denoted by $f = 1, \dots, F$. For a given time period ($t = 1, \dots, T$) each fund is associated with a risk-adjusted performance metric, $RAP_{ft} \equiv \hat{\alpha}_{ft}$ (e.g. 'Alpha'), that classifies each fund in one of the n non-overlapping segments (states) of the resulting ranking distribution of RAP_{ft} across all funds. In the following period ($t + 1$) the RAP_{ft+1} metrics are calculated again generating a new ranking distribution. One can then represent the ranking dynamics of mutual funds' performance using a Markov Chain. The Markov Chain is determined by the transition matrix, whose typical element p_{ij} denotes the transition probability that a fund is in state j at time $t + 1$ given it was placed in state i in time t, as given below:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \tag{2}$$

where the following restrictions hold: *(i)* $p_{ij} \geq 0, \forall i, j$, i.e. all probabilities are non-negative, and *(ii)* $\sum_{i=0}^{n} p_{ij} = 1, \forall i$, the set of states is exhaustive. The matrix $\mathbf{P}$ summarizes the n^2 transition probabilities that correspond to all possible movements while the evolution of the cross-sectional distribution ($\boldsymbol{d}$) over time is described by as below:[5]

$$\boldsymbol{d}_t = \boldsymbol{d}_{t-1}\mathbf{P} \tag{3}$$

---

[5] It might well be the case that transition probabilities are not constant over time hence the Markov Chain is non-stationary. In this case, the transition matrix from period t to period $t + 1$ is $\mathbf{P}(t)$ with its typical element being $p_{ij}^{(t)}$, to denote the time variation in transition probabilities. This is in fact what we find in our data, but for the sake of exposition in text we focus on stationary Markov chains.

Recall that one of the approaches in assessing performance persistence is to classify a fund as a 'winner' ('loser') when its risk-adjusted performance is above (below) the median of the empirical cross-sectional distribution of fund performance (e.g. Goetzmann and Ibbotson, 1994; Brown and Goetzmann, 1995; Malkiel, 1995; Elton et al., 1996). It is then easy to track its dynamics using a 2×2 contingency (transition) matrix:

$$\begin{bmatrix} N_{WW} & N_{WL} \\ N_{LW} & N_{LL} \end{bmatrix} \tag{4}$$

where N_{ij}, $i = W, L$ and $j = W, L$, denotes the counts of transitions from initial ranking i to post-ranking j. One can then examine whether persistence by means of the cross product ratio.6 Our suggestion is to generalize this setting, and track transition probabilities from ranking i to ranking j, for any initial classification. As we explain below these are analogous to counting transitions from state i to state j.

If the mutual fund market exhibits no persistence, the probabilities of remaining in a particular ranking (state) or transiting to any other ranking in period $(t + 1)$ would be equal to each other, and all of them would equal $(1/n)$. Hence, movements would be generated from a uniform distribution and the $n \times n$ no persistence-compatible transition matrix $\mathbf{P}^{NP}$ should take the following form:

$$\mathbf{P}^{NP} = \begin{bmatrix} 1/n & 1/n & ... & 1/n \\ 1/n & 1/n & ... & 1/n \\ \vdots & \vdots & & \vdots \\ 1/n & 1/n & \cdots & 1/n \end{bmatrix} \tag{5}$$

Consider for instance the 'winner-loser' case, employed in the early literature on performance persistence. During the ranking period, all funds by definition have a probability (1/2) to be above or below the median benchmark that defines performance, hence a Bernoulli distribution. In the evaluation period, under the null of no persistence, a fund classified as 'winner' in the ranking period, again faces a Bernoulli distribution with probability (1/2) of being classified either as a winner or as a looser. Similar results hold for the 'loser' case.7 Hence the transition matrix should take values $p_{ij} = 1/2$ for $i, j = W, L$. In the general case,

---

6 Under the null of "no-persistence" the odds/cross-product ratio $CCR = (N_{WW} \times N_{LL})/(N_{WL} \times N_{LW})$ should be unity (Brown and Goetzmann, 1995). The test statistic under the assumption of normally that assumes the following form (Chrinstesen, 1990): $Z = \ln(CCR)/\sigma_{\ln(CCR)}$, where $\sigma_{\ln(CCR)} = \sqrt{N_{WW}^{-1} + N_{WL}^{-1} + N_{LW}^{-1} + N_{LL}^{-1}}$.

7 See Agarwal and Naik (2000) and Brown et al (1999) for an application in hedge funds performance persistence.

where funds are ranked into n initial categories, it follows that the probability that a fund starts in rank i and ends up in rank j during the evaluation period, is just a multinomial distribution with probabilities $p_{ij} = 1/n$.

It then becomes apparent that the elements along the main diagonal of the empirical transition matrix **P** contain important information, as they measure the probabilities that a fund remains in its initial group ranking; or equivalently the decreased probability of changing states (moving up or down the ranking distribution). Thus, finding that the elements of the main diagonal of **P** have significantly deviated from the main diagonal of the (*no-persistence*) benchmark matrix $\mathbf{P}^{NP}$, would indicate the presence of persistence.

Persistence will be absent if the empirical transition matrix satisfies the following conditions:

 – *The Markov Chain is irreducible,* implying that all transition probabilities are non-zero, thereby every state of the process may be visited from any given initial state, with positive probability. Hence irreducibility requires that $p_{ij} > 0, \forall\, i,j$.

 – *Transition probabilities are uniform,* so that all states have equal probabilities of being visited. Hence uniformity requires: $p_{ij} = 1/n, \forall\, i,j$.

2.3 Assessing Persistence using Mobility Indices

We assess the degree of performance persistence using the empirical transition matrices (obtained from a Markov model for a mutual fund performance metric), focusing on the properties of the matrix *per se*; and by comparing the empirical transition matrix with appropriately selected 'theoretical' transition matrices. Several mobility indices previously applied in exploring income mobility (Shorrocks, 1978; Geweke, *et al*., 1986; Bigard *et al*, 1998) and credit rating migration (Jafry and Schuermann, 2004; Trueck and Rachev, 2006) are used to assess the degree of movement exhibited in the empirical transition matrices. Moreover, we also make comparisons of mobility over time, by calculating the indices for different periods.

Initially, following Bigard *et al*. (1998) we first compute the immobility ratio (***IR***), which records the *percentage* of funds which stay in the *same relative performance ranking* during the post-ranking period, hence capturing the persistence in mutual fund relative performance.

We then employ a different set of mobility indices due to Prais (1955), Shorrocks (1978) and Sommers and Conlisk (1979) i.e. M_{PS}, M_S, and M_{SC}, respectively, all of which are indices based on the eigenvalues of the estimated transition matrices. In particular, letting λ_i denote the i-th eigenvalue of the estimated transition matrix, these indices are defined as:

$$M_{PS} = \frac{1}{n-1}\left(n - tr(\mathbf{P})\right)$$

$$M_S = 1 - |\det(\mathbf{P})| = 1 - \left|\prod_{i=1}^{n} \lambda_i\right|$$

$$M_{SC} = 1 - |\lambda_2|$$

where $tr()$ and $\det()$ denote the trace and the determinant of the transition matrix. These indices range between zero and unity, where zero denotes complete persistence and unity denotes complete lack of persistence in performance rankings. Another relevant indicator we employ is *half-life* ($h = -\log(2)/\log(\lambda_2)$), which shows the speed of convergence towards the equilibrium (unconditional) distribution of performance rankings. In other words, *half-life* measures how long it takes for the current (ranking period) performance ranking to cover half of the deviation from the long-run performance ranking (Theil, 1972). This indicator ranges between zero and infinity corresponding to perfect mobility and total immobility respectively.

Persistence may also be assessed by comparing the properties of the empirical transition matrices to the properties of the benchmark of complete inertia/persistence (i.e. the identity matrix). These comparisons can be based on the Euclidean distance of the transition matrices, employing indices such as the D_{L^1} (Israel *et al.*, 2001), D_{L^2} (Banglia *et al.*, 2002) and $D_{L^{max}}$ (Trueck, 2004). Alternatively one may resort to singular-values-based indices (M_{SVD} and D_{SVD}) proposed by Jafry and Schuermann (2004), Trueck (2004) and Trueck and Rachev (2006) who built on the idea that a transition matrix expresses quantitatively how a given state-vector will move (migrate) from one epoch (ranking) to the next. In this sense, these metrics measure how relative rankings have changed

The Euclidean distance measures, as well as the M_{SVD} and D_{SVD} indices compare the empirical transition matrix to an appropriately selected benchmark matrix, typically the identity matrix (complete persistence). Similarly, we may implement a version of these indices that is directly linked to the notion of performance persistence. This is simply done by replacing the identity matrix with $\mathbf{P}^{NP}$ (see (6)), whose properties are determined by the assumption of no persistence. In this vein, the indices we propose capture how close the observed transition probabilities are to the *no persistence* benchmark. These are simply defined as follows:[8]

---

[8] In our case where $n = 4$, the first two indices (D_{L^1} and D_{L^2}) take values that range between zero (no distance from benchmark) and eight (maximum distance from benchmark). On the other hand, the

$$D_{L^1}(\mathbf{P}, \mathbf{P}^{NP}) = \sum_{i=1}^{n} \sum_{j=1}^{n} |p_{ij} - p_{ij}^{NP}| \in [0, 2n] \tag{6}$$

$$D_{L^2}(\mathbf{P}, \mathbf{P}^{NP}) = \sum_{i=1}^{n} \sum_{j=1}^{n} (p_{ij} - p_{ij}^{NP})^2 \in [0, 2n] \tag{7}$$

$$D_{L^{max}}(\mathbf{P}, \mathbf{P}^{NP}) = max_{i,j} |p_{ij} - p_{ij}^{NP}| \in [0, 1] \tag{8}$$

Similarly, to construct new singular-values-based indices, we calculate the 'immobility' matrix:

$$\mathbf{P}^{IM} = \mathbf{P} - \mathbf{P}^{NP} \tag{9}$$

Then, following Jafry and Schuermann (2004), we use the average of all singular values of $\mathbf{P}^{IM}$:

$$M_{SVD}^{IM} = \frac{1}{n} \sum_{i=0}^{n} \sqrt{\lambda_{IM,i}[(\mathbf{P}^{IM})'(\mathbf{P}^{IM})]} \tag{10}$$

Furthermore, in line with Trueck (2004) and Trueck and Rachev (2006), we measure the distance between the empirical transition matrix $\mathbf{P}$ and $\mathbf{P}^{NP}$ by calculating the metric:

$$D_{SVD}(\mathbf{P}, \mathbf{P}^{NP}) = \frac{1}{n} \sum_{i=1}^{n} \sqrt{\lambda_{\mathbf{P},i}(\mathbf{P}'\mathbf{P})} - \frac{1}{n} \sum_{i=1}^{n} \sqrt{\lambda_{\mathbf{P}^{NP},i}(\mathbf{P}^{NP'}\mathbf{P}^{NP})} \tag{11}$$

In principle one may use any of these indices to assess mobility. In our empirical analysis we employ all of them, however it is fruitful to offer the interested reader a critical appraisal of their advantages and disadvantages.

A major shortcoming of "eigenvalues-based" and "cell-by-cell" indices is that they remain agnostic to variations in the distribution of the off-diagonal elements, *i.e.*, are not "distribution discriminatory" (Jafry and Schuermann, 2004). In other words, they do not distinguish between transition matrices with a different configuration of transition probabilities in the off-diagonal cells. In contrast, "singular values-based" indices are "distribution discriminatory", a property that allows them to discern differences in the off-

---

other three distance metrics we employ ($D_{L^{max}}$, M_{SVD}, and D_{SVD}) take values between zero (no distance) and one (maximum distance).

diagonal elements. Knowledge of this configuration is crucial since distant transitions compared to close ones clearly have different implications for persistence.

3 Data Issues and Background Analysis

3.1 Data

Data on Mid-Capitalization US Equity mutual funds are obtained for the period from January 1982 to December 2005 from the Center for Research in Security Prices (CRSP). We note that the number of active funds has increased dramatically during the period under scrutiny starting from 136 funds in January 1982 and soaring to 2590 in December 2005. A time plot of the number of active funds results in an S-shaped curve as shown in Figure 1. One immediately observes the rapid acceleration that took place during this period. Gruber (1996) has pointed out this phenomenon and called it a puzzle given the fact that the average performance of funds is typically inferior to that of the market.

Figure 2 depicts the demographics of the particular mutual fund category in the US. Annual entry and exit rates have been calculated as the percentage of new funds and funds exiting to the total number of funds, respectively. The data show that during the period 1982-1987 there were only entries and no exits whatsoever. In the remaining sample period exits occurred, although entries continued to dominate, exceeding exits on average by 6:1.

The start-ups (entries) and deaths (exits) that result in a time-varying population of the fund market is of course a natural phenomenon, which however may adversely affect the econometric analysis used to assess performance persistence. Of particular importance is the exit of funds, which among other things may reflect inferior performance, and as such would be related to the issue at hand. We have investigated various properties of the exit rate in order to evaluate its significance for our study. We found that the annual mean exit rate has been 0.24 percent, which represents a minute proportion of the population and therefore conclude that it does not pose a serious threat to the assessment of performance persistence that will follow (see also Table B.2 in the appendix for detailed summary statistics of the exit and entry rates).

The sample we employ does not suffer from survivorship bias as identified in Brown *et al.* (1992) and Brown and Goetzmann (1995), where only funds that exist at the end of the sample are included in the analysis. Instead, we use information from all funds available both during ranking, and post-ranking periods, that is both in the classification and evaluation periods we use all funds alive, but none of our findings is sensitive to this choice.[9]

---

[9] In previous work, we have also experimented in obtaining the quartiles using the same funds for the classification and the evaluation periods. Our results are virtually unchanged quantitatively and quantitatively.

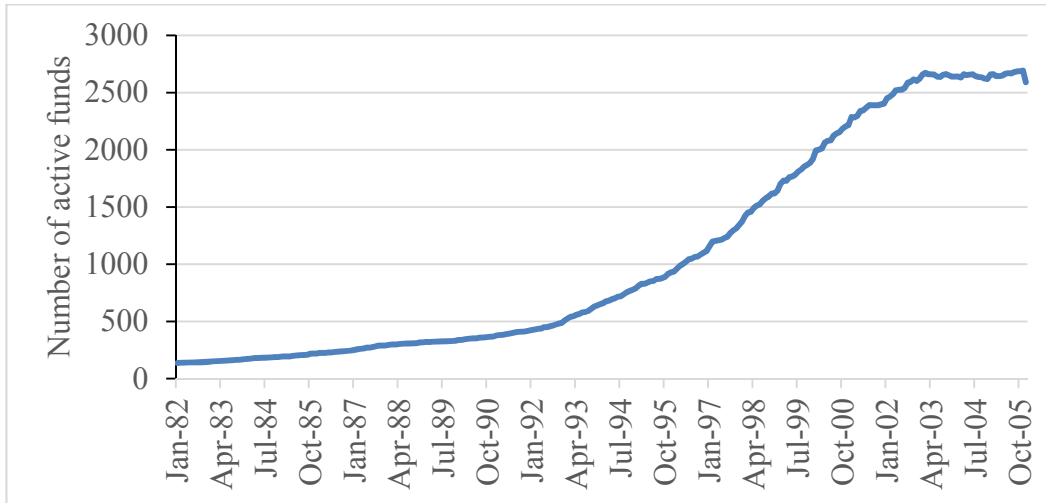Figure 1. Time Evolution of Number of Active Funds

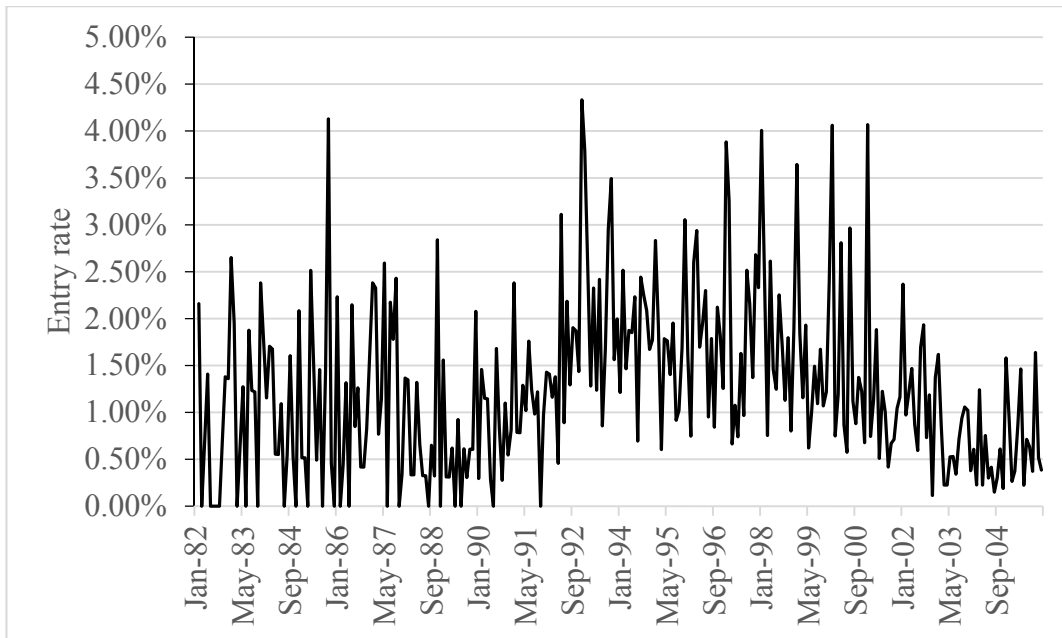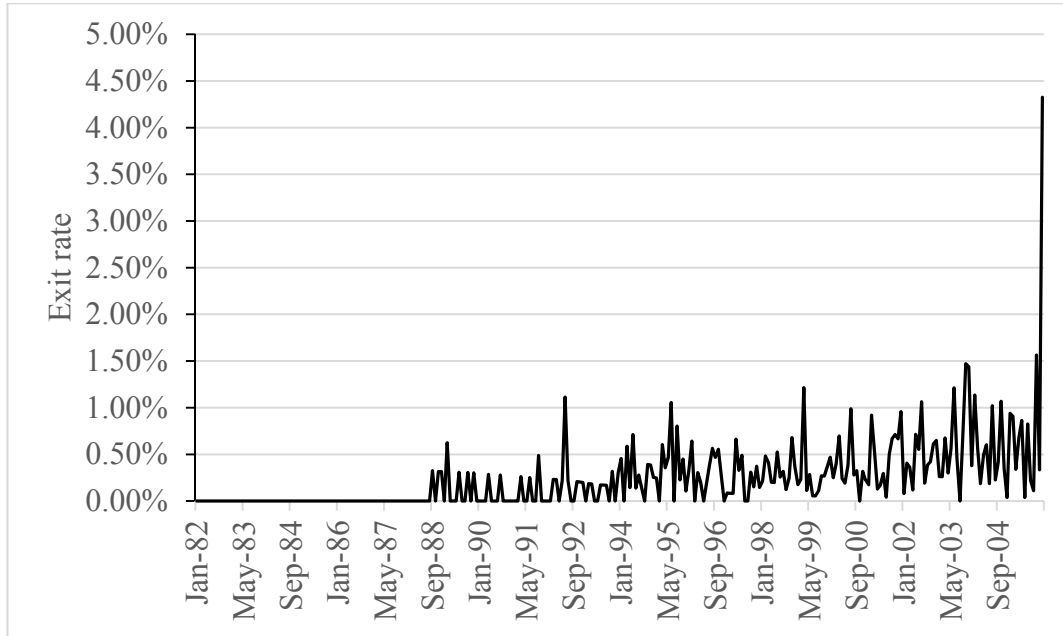Figure 2. Time Evolution of Entry and Exit Rates: Entry rate

Figure 2. Time Evolution of Entry and Exit Rates: Exit rate

Note: The figure displays the entry and exit rates (as a percentage of total population of funds) respectively for each year in the sample.

3.2 Performance Metrics and Performance Persistence

In order to obtain our risk-adjusted performance metrics, we divide the period 1982-2005 into 8 non-overlapping three-year intervals and use monthly data to estimate the parameters of the four-factor model to recover each fund's abnormal return. We chose a three-year investment horizon. Focusing on a shorter horizon (say 24 months) would severely reduce the precision of our estimates.[10] We should highlight here that we do retain in our analysis any fund for which we have at least 18 monthly observations either in the classification or in the evaluation period. On the other hand, focusing on longer investment horizons would cloud the assessment of persistence *per se*, since previous studies have shown that any persistence, if present, tends to disappear in longer horizons. Table 1 reports summary statistics for the estimated Alphas for each three-year interval.

---

[10] In an earlier draft we experimented with investment horizons of 24 months, which resulted in less precisely estimated parameters and risk-adjusted returns. However, the main results of our analysis are not sensitive to this choice.

Table 1. Descriptive Statistics for Alphas by 3-Year Interval

Time Interval	Mean	Std. Dev.	Min	Max	N
1982-1984	0.0036	0.44	-1.69	1.19	**155**
1985-1987	0.71	0.38	-1.06	2.48	**224**
1988-1990	1.01	0.5	-1.41	2.51	**317**
1991-1993	-0.068	0.45	-2.73	4.57	**412**
1994-1996	-0.12	0.37	-4.6	1.63	**787**
1997-1999	-0.16	0.49	-2.66	3.43	**1392**
2000-2002	-0.13	0.62	-6.33	3.68	**2168**
2003-2005	0.16	0.25	-2.29	1.46	**2501**

Note: Alphas denote risk-adjusted excess returns from a four factor selection model

$$R_{ft} = \alpha_{ft} + \beta_{1f}RMRF_t + \beta_{2f}SMB_t + \beta_{3f}HML_t + \beta_{4f}MOM_t + u_{ft}$$

which are expressed as percentages. *N* denotes the month-fund data points in the particular 3-year period.

Having obtained abnormal returns, we sort funds according to this measure. For each sorting, we form quartiles, and we then track how the funds in each quartile perform in the following period.

Table 2 reports the average abnormal return of funds in each quartile (equally weighted average across funds) during the ranking and post-ranking (evaluation) periods. Note first, that there is ample evidence of performance persistence, especially until the mid-1990s. In particular, we note that there is a tendency for repeat performers, for the first two ranking/post-ranking periods (where all funds retain their relative ranking). Furthermore, in the following two ranking/evaluation periods, we observe that funds ranked in the top quartile tend to outperform other funds, while similarly funds classified as the worst performers, tend to remain in the last quartile in the post-ranking period. In contrast, during the next two ranking/evaluation periods, we find that only performers in the top quartiles tend to repeat their rankings (with only the top 25% performers during 1997-1999/2000-2002). Finally, during the last ranking/evaluation period, there is virtually no tendency for performance persistence.

In order to examine this issue in a more formal quantitative manner, we calculate Spearman rank-correlations, which allow the examination of similar rankings during the ranking and evaluation period. We find that initial rankings tend to be repeated during the evaluation period, and this holds until the late 1990s. Instead, we also find that there is no tendency for ranking repetition during the 2000s. In other words, the existence of any performance persistence shows a tendency to vanish towards the end of our sample. We

Table 2. Risk-adjusted performance quartiles: Ranking and Evaluation Period

Quartile	Ranking 1981-1984	Evaluation 1985-1987	Ranking 1985-1987	Evaluation 1988-1990	Ranking 1988-1990	Evaluation 1991-1993	Ranking 1991-1993	Evaluation 1994-1996
1	0.526	0.895	1.165	1.104	1.580	0.024	0.384	-0.033
2	0.138	0.730	0.797	1.055	1.163	-0.066	0.016	-0.176
3	-0.092	0.702	0.619	1.014	0.932	-0.060	-0.151	-0.084
4	-0.553	0.474	0.276	0.967	0.367	-0.244	-0.522	-0.277
Average	0.004	0.700	0.714	1.035	1.010	-0.062	-0.068	-0.135
Rho	0.432**		0.096		0.223**		0.230**	
t-stat	5.928		1.438		3.941		4.500	
N	155		224		317		412	

Quartile	Ranking 1994-1996	Evaluation 1997-1999	Ranking 1997-1999	Evaluation 2000-2002	Ranking 2000-2002	Evaluation 2003-2005
1	0.291	-0.171	0.422	-0.028	0.661	0.132
2	-0.027	-0.191	-0.083	-0.171	-0.066	0.157
3	-0.212	-0.262	-0.302	-0.156	-0.343	0.160
4	-0.558	-0.226	-0.684	-0.061	-0.796	0.150
Average	-0.127	-0.210	-0.162	-0.104	-0.136	0.150
Rho	0.089**		-0.004		-0.018	
t-stat	2.352		-0.148		-0.738	
N	787		1392		2168	

Notes: The table lists average three-year estimates during the ranking and evaluation (post-ranking) periods for quartiles of funds sorted according to the performance estimates during the ranking period. We base our rankings risk-adjusted measures of performance, α_{ft}, estimate*d* from the four factor model:

$$R_{ft} = \alpha_{ft} + \beta_{1f}RMRF_t + \beta_{2f}SMB_t + \beta_{3f}HML_t + \beta_{4f}MOM_t + u_{ft}.$$

The table also reports Spearman rank correlation coefficients (rho) as well a *t*-test statistic for no correlation in repeated rankings (during the ranking and the post-ranking period). The sample for each ranking/evaluation period is given in the last row of the table. One and two asterisks indicate two-tailed significance at the 5% and 1% levels respectively.

should stress here, that had we followed the root of averaging performance metrics both across funds and across time, this would have cloud our last finding that there seems to be no performance persistence towards the end of the sample.

Our evidence thus far show that the behavior of mid-cap US Equity mutual funds has been evolving over time, from a situation where there were indeed repeat performers during the early 1980s and until the mid-1990s, to a situation where initial rankings contain no information about future rankings, during the 2000s. We explore this change further in what

follows, by means of a Markov Chain model, which allows us to assess how funds have transited from any initial performance ranking to their subsequent rankings.

4 Main Results

4.1 Estimation and Testing Transition Probabilities

For each three-year interval we rank funds according to their estimated Alphas and classify them into a specific quartile. In the subsequent time interval (the evaluation period) we re-rank funds and calculate all transition counts. Then employing Anderson and Goodman's (1957) maximum likelihood estimator we estimate the transition probabilities, with our results summarized in Table 3. This procedure leads to the estimation of seven empirical transition matrices. Having estimated the empirical transition matrices we test hypotheses regarding specific values for the transition probabilities following Cochran (1952) and Anderson and Goodman (1957).[11]

Before proceeding, two important issues related to our analysis deserve special attention. The *first*, relates to the treatment of entry and exit in our sample. In practice 'defunct' funds are used to obtain the sort and quartiles during the ranking period. Instead, new funds are included in the ranking period – provided they have a sufficient number of observations to estimate risk-adjusted abnormal returns – or else, they are included for sorting and forming quartiles in the post-ranking period. We should also point out that although we do not explore this route here, we could include another 'liquidation' state in our analysis, to capture exit. Note however that, that 'liquidation' is an *absorbing* state: eventually all funds alive (in the limit) have to liquidate. Although theoretically possible, such a 'liquidation' state in never observed for all the funds in our sample, as during each and every period there are funds which are active. We feel that the 'liquidation' state is – from a practical viewpoint – is rather uninteresting.[12] *Second*, we find that the estimated transition probabilities do not remain constant over time; hence the assumption of stationarity is violated (see Table B.3 in the appendix). However, the results we present (mobility indices) do not depend in any way on the assumption of stationarity. Instead, this finding is desirable, to the extent that it allows us to investigate how relative performance rankings may have changed over time.

Having estimated the transition matrices (Table 3), we first note that the elements along the main diagonal tend to be larger than the rest of the elements in each row, signifying that other and consequently irreducibility of the Markov Chain is satisfied. To understand this

---

[11] Formal discussion of these estimation and testing issues is provided in Appendix B.

[12] We would like to thank an anonymous referee for raising this issue.

Table 3. Estimated Transition Matrices

$$P(8587) = \begin{pmatrix} 0.41 & 0.36 & 0.15 & 0.08 \\ 0.25 & 0.28 & 0.28 & 0.20 \\ 0.19 & 0.32 & 0.35 & 0.14 \\ 0.10 & 0.10 & 0.33 & 0.46 \end{pmatrix} \qquad P(8890) = \begin{pmatrix} 0.31 & 0.31 & 0.15 & 0.24 \\ 0.23 & 0.21 & 0.36 & 0.20 \\ 0.16 & 0.36 & 0.31 & 0.16 \\ 0.31 & 0.16 & 0.22 & 0.31 \end{pmatrix}$$

$$P(9193) = \begin{pmatrix} 0.36 & 0.28 & 0.21 & 0.15 \\ 0.22 & 0.33 & 0.22 & 0.22 \\ 0.36 & 0.16 & 0.21 & 0.27 \\ 0.14 & 0.23 & 0.27 & 0.37 \end{pmatrix} \qquad P(9496) = \begin{pmatrix} 0.34 & 0.30 & 0.22 & 0.15 \\ 0.21 & 0.29 & 0.24 & 0.27 \\ 0.24 & 0.26 & 0.31 & 0.19 \\ 0.13 & 0.22 & 0.31 & 0.34 \end{pmatrix}$$

$$P(9799) = \begin{pmatrix} 0.28 & 0.24 & 0.23 & 0.25 \\ 0.20 & 0.32 & 0.25 & 0.23 \\ 0.14 & 0.29 & 0.27 & 0.30 \\ 0.25 & 0.20 & 0.27 & 0.28 \end{pmatrix} \qquad P(0002) = \begin{pmatrix} 0.30 & 0.25 & 0.25 & 0.21 \\ 0.19 & 0.31 & 0.26 & 0.24 \\ 0.22 & 0.26 & 0.31 & 0.21 \\ 0.38 & 0.20 & 0.17 & 0.25 \end{pmatrix}$$

$$P(0305) = \begin{pmatrix} 0.29 & 0.16 & 0.21 & 0.33 \\ 0.21 & 0.25 & 0.29 & 0.25 \\ 0.21 & 0.27 & 0.32 & 0.20 \\ 0.24 & 0.27 & 0.22 & 0.27 \end{pmatrix}$$

Notes: Table reports the estimated transition matrices for each sub-sample period using Maximum Likelihood. Each element of each transition matrix p_{ij} corresponds to the estimated probability of transiting from state i to state j, across states ($n = 4$). The first state identifies the funds who are in the top 25 percent of performers, the second state the funds between 26 and 50 percent of performers, the third state funds belong between 51 and 75 percent of performers, and finally the fourth state (76 to 100 percent) shows the worst performers.

finding, consider a fund being ranked in the top performers during the classification period. There is a non-zero probability that during the evaluation period, it will end up either in the same ranking or it will change its ranking, even ending up in the worst performers group. Similarly, any fund being initially in the worst performers, may end up in any possible classification during the evaluation period.

This finding implies that a mutual fund whose risk-adjusted performance in a given three-year interval places it in a particular ranking bracket may transit to any other bracket with a non-zero probability.

Table 4. Uniformity tests $H_0: p_{ij} = \frac{1}{4}, \forall i, j$

	P(85-87)	P(88-90)	P(91-93)	P(94-96)	P(97-99)	P(00-02)	P(03-05)
LR-test	34.04	18.11	25.03	22.12	21.47	47.22	57.63
	[0.00]	[0.11]	[0.01]	[0.03]	[0.02]	[0.00]	[0.00]
Pearson-test	32.25	17.84	24.92	21.02	20.15	48.56	57.31
	[0.00]	[0.12]	[0.01]	[0.05]	[0.04]	[0.00]	[0.00]

Note: Table reports likelihood ratio and Pearson-type tests of the hypothesis of uniformity for all estimated transition matrices. *P*-values in square brackets. Both tests are distributed as chi-squared variates with 12 d.f.

We next explore whether the property of uniformity is satisfied, with results reported in Table 4. Uniformity is a sufficient condition for the absence of persistence and in our setup takes the form of equality of each transition probability to (1/4). Finding that this restriction is supported by the data, would imply that not only all states communicate with each other, but also the probabilities of transiting from a given state to any other would be equal. Our findings show that in six out of seven cases the hypothesis of uniformity is rejected, while only for the time-interval of 1988-1990 the hypothesis is marginally supported by the data (*p*-value, 0.11). It becomes evident that for the period 1982-2005 perfect mobility in its extreme form (*i.e.*, uniformity) is rejected. Nevertheless, since irreducibility is satisfied, we may infer that the finding of persistence is driven by non-uniformity. Thus, all states (performance brackets) communicate and therefore funds may transit to any state (*i.e.*, irreducibility), albeit with unequal probabilities.

4.2 Evaluating Persistence by Means of Mobility Indices

Our findings thus far show that transition probabilities are not uniform; hence not all states are equally likely to be visited, starting from any initial state. In addition, we have documented that transition probabilities are indeed non-zero, but they are also time-varying: they change depending on the particular raking and evaluation period examined. Given these findings, in order to explore further the existence and evolution of persistence, we proceed by calculating various mobility indices described in Section 2. Our results are summarized in Table 5.

Starting with the general mobility indices, namely *IR* (the immobility ratio), *MU* (the percentage of funds that improve their relative ranking) and *MD* (the percentage of funds that worsen their relative ranking), we first compare them amongst themselves, since by construction they add up to 100 percent at any given period. Note that under the assumption of perfect mobility each of these indices should equal 33.3 percent (*i.e.* equal probability of staying put or transiting upwards or downwards). We find that in no ranking/evaluation

Table 5. Mobility Indices

Index	P(85-87)	P(88-90)	P(91-93)	P(94-96)	P(97-99)	P(00-02)	P(03-05)
	General Mobility Measures						
IR	37.45%	28.45%	31.57%	31.77%	28.90%	29.15%	28.13%
MU	30.00%	35.20%	34.00%	33.88%	37.44%	35.48%	36.28%
MD	32.55%	36.26%	34.42%	34.35%	33.66%	35.37%	35.59%
	Eigenvalues based						
M_{PS}	0.8340	0.9528	0.9124	0.9098	0.9480	0.9447	0.9583
M_{SC}	0.6561	0.7829	0.7918	0.8516	0.9050	0.9411	0.8740
M_S	0.9998	0.9987	0.9993	0.9996	0.9998	0.9997	0.9999
h (in months)	0.6493 (23 months)	0.4538 (16 months)	0.4417 (16 months)	0.3634 (13 months)	0.2945 (11 months)	0.2448 (9 months)	0.3346 (12 months)
	Cell by cell based						
$D_{L^1}(\mathbf{P}, \mathbf{I})$	5.004	5.717	5.474	5.459	5.688	5.668	5.750
$D_{L^2}(\mathbf{P}, \mathbf{I})$	2.212	2.798	2.558	2.518	2.718	2.711	2.782
$D_{L^{max}}(\mathbf{P}, \mathbf{I})$	0.725	0.786	0.792	0.711	0.727	0.746	0.753
$D_{L^1}(\mathbf{P}, \mathbf{P}^{NP})$	1.5796	1.0325	0.9782	0.8223	0.5296	0.6199	0.5938
$D_{L^2}(\mathbf{P}, \mathbf{P}^{NP})$	0.4565	0.2843	0.2894	0.2435	0.1722	0.2061	0.1778
$D_{L^{max}}(\mathbf{P}, \mathbf{P}^{NP})$	0.2115	0.1136	0.1162	0.1201	0.1106	0.1288	0.0851
	Singular values based						
$D_{SVD}(\mathbf{P}, \mathbf{I})$	0.6311	0.7158	0.6865	0.6846	0.7123	0.7113	0.7198
$D_{SVD}(\mathbf{P}, \mathbf{P}^{NP})$	0.1461	0.1061	0.1033	0.0801	0.0568	0.0686	0.0584

Note: **P** denotes the empirical transition matrix. **I** and **P**^{NP} stand for the identity matrix and the a square matrix with all its elements equal to 1/4.

period such a configuration is encountered. Moreover, in every time-period with the exception of 1985-1987 both *MU* and *MD* are always greater than *IR*, showing a decaying rate of inertia and therefore increasing movements in either direction. Although both *MU* and *MD* increase over time, their difference is small providing evidence for almost equal bi-directional mobility. The overall conclusion reached is that the mutual fund market is characterized by a considerable degree of mobility, which evolves over time towards higher levels of mobility.

We next turn to the "eigenvalues-based" mobility indices M_{PS}, M_{SC}, and M_S whose domain is [0, 1], and find that they attain different values. In particular we find that M_{PS} and M_{SC} increase over time, growing ever closer to the value of unity, which corresponds to the case of perfect mobility, while M_S is always close to one. This taken at face value suggests that performance persistence is becoming less and less important. On the other hand, these

measures cannot differentiate between alternative configurations in the off-diagonal elements, and consequently do not provide any information with respect to the distribution of probabilities across neighboring and distant states. In any case, they suggest that the mutual fund market is characterized by a high degree of mobility.

The estimated half-life metric reveals that the time needed to cover half of the distance from the unconditional distribution (long-run ranking) is considerably reduced during the 1990's. For instance, when translated into months the indicator shows that in 1985-1987 it would take up to 23 months to cover 50 percent of disequilibrium, whereas during 2003-2005 it would only take 12 months. Again, we interpret this as evidence in favor of increase mobility, hence of reduced performance persistence over time.

Last, we examine the "cell-by-cell based" (Euclidean distance measures) and "Singular values-based" indices. Recall that these are defined relative to the identity matrix, a benchmark of complete immobility (complete persistence) and a benchmark of perfect mobility (no persistence). We first find that when calculated using the identity matrix $(D_{L^1}(\mathbf{P}, \mathbf{I}), D_{L^2}(\mathbf{P}, \mathbf{I}), D_{L^{max}}(\mathbf{P}, \mathbf{I})$ and $D_{SVD}(\mathbf{P}, \mathbf{I}))$ as the comparison benchmark, the indices attain considerably higher values than when the distance is measured from the perfect mobility benchmark $((D_{L^1}(\mathbf{P}, \mathbf{P}^{NP}), D_{L^2}(\mathbf{P}, \mathbf{P}^{NP}), D_{L^{max}}(\mathbf{P}, \mathbf{P}^{NP})$ and $D_{SVD}(\mathbf{P}, \mathbf{P}^{NP}))$. This provides solid evidence that the empirical transition matrix in all cases is 'closer' to the perfect mobility benchmark, rather than the perfect immobility benchmark.

Useful conclusions can also be derived from the time evolution of the relevant indices, which are depicted in Figure 3 on the next page. The dynamic behavior of all Euclidean and Singular values-based indices reveals an increasing similarity of the estimated transition matrices to the perfect mobility benchmark matrix. The distance indices using the identity matrix (perfect immobility) as the reference point tend to increase over time, while they systematically decrease when the perfect mobility (no persistence matrix) is used as the reference point. Thus, based on this evidence we document that the degree of mobility has unequivocally increased.

4.3 Discussion

Our findings are partly in contrast to the existing evidence on mutual fund performance persistence. For instance, Bollen and Busse (2004) document that performance persistence is extremely short-lived, well below the one year horizon. Instead, we find that performance persists even at longer horizons, of about three years. One possible source of difference is that we use a different method to sort funds and measure persistence. In particular, although we sort by prior three-year risk-adjusted performance and we measure post-ranking performance by abnormal return in the following three years – Bollen and Busse (2004) do the same type
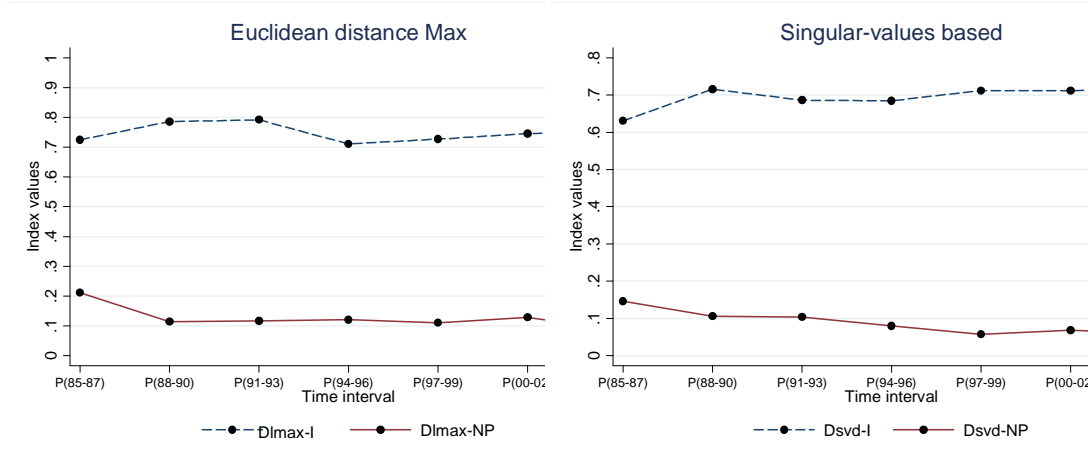
Figure 3. Euclidean and Singular Values Indices

3.A: $D_{L^1}(\mathbf{P},\mathbf{I})$, $D_{L^1}(\mathbf{P},\mathbf{P}^{\mathbf{NP}})$ **3.B:** $D_{L^2}(\mathbf{P},\mathbf{I})$, $D_{L^2}(\mathbf{P},\mathbf{P}^{\mathbf{NP}})$
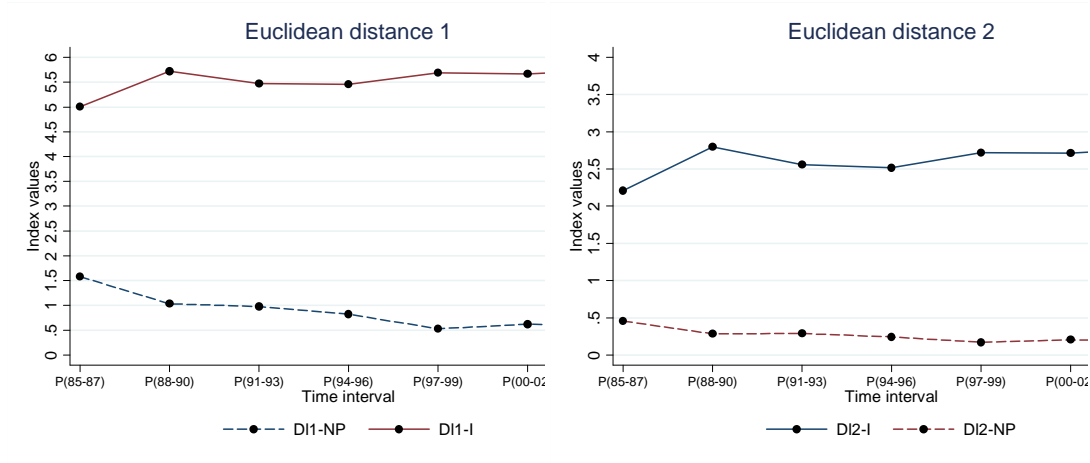
3.C: $D_{L^{\max}}(\mathbf{P},\mathbf{I})$, $D_{L^{\max}}(\mathbf{P},\mathbf{P}^{\mathbf{NP}})$ **3.D:** $D_{SVD}(\mathbf{P},\mathbf{I})$, $D_{SVD}(\mathbf{P},\mathbf{P}^{\mathbf{NP}})$

of sorting and evaluation using quarters – we do not time-average our results, allowing for much more variation and richer dynamics. Furthermore, Bollen and Busse (2004) focus on a fixed, time-invariant sample of mutual funds not taking into account entry and exit, whereas by also track new funds and liquidated funds in our analysis.

Our findings are in some respects in line with those reported in Carhart (1997), but are based on a slightly different methodology which also provides new insights. In particular, Carhart reports a monthly top-decile four-factor α_f of 0.02%, when sorting by prior year risk-adjusted return whereas we find that the top-quartile four-factor α_f is about 0.05% per month.

Furthermore, Carhart finds no evidence of persistent superior ability after controlling for the momentum anomaly, while we find that performance persists, even after controlling for momentum. In our analysis, we measure post-ranking performance by abnormal return in the following three years. In contrast, Carhart estimates post-ranking performance by a concatenated series of post-ranking returns, which does not allow for time variation in either risk-adjusted returns or even in factor risk loadings. Instead, since we re-estimate the four-factor model during the ranking and post-ranking periods, we allow for time variation in both risk-adjusted excess returns and factor loadings, and we also document that performance persists until the late 1990s.

We have also shown that performance persistence vanishes only towards the end of the sample, during the 2000s, while being detected over long horizons (three-year intervals). One could argue that performance is a short-term phenomenon generated by (short-term) informational advantages which some managers are able to exploit, we instead argue that this is probably a good description for periods during which competition is not strong enough in the market, while when the market has expanded enough – both in terms of participants and alternative portfolios - these advantages tend to dissipate. In this vein, our findings is in line with the model of Berk and Green (2004), where the short-term performance could be the outcome of actions by investors who rely on a group of fund managers with different levels of ability. In this context, rational investors form beliefs about ability based on past performance, and allocate their funds to "more able" managers. But this abnormal performance eventually vanishes due to its decreasing returns to scale and/or due to the possibility that a manager increases fees. We do find that such a reduction in persistence takes place, as the market becomes more competitive, with a growing number of competitors in the market. This finding is also in contrast to those in Bollen and Busse (2004) who track a fixed number of funds over time

5. Conclusions

In this paper, we revisit the issue of persistence in mutual fund performance. By modeling the dynamics of fund performance as a Markov Chain, we are able to focus on the mobility manifested in successive performance ranking of US funds. We rank funds every three years according to risk-adjusted returns, and then measure the risk-adjusted return of quartiles of funds over the following three-year period. We find that until the late 1990s, there is a tendency for repeat performers in the post ranking periods. We then model the transitions from an initial performance ranking to a subsequent ranking as a Markov Chain, estimating the relevant transition probabilities. In this context, we also assess the extent of performance persistence by means of several mobility indices. According to our findings for the period 1982-2005 persistence is present. Nevertheless, given that irreducibility is satisfied we conclude that funds may transit to any state albeit with unequal probabilities.

The overall conclusion reached is that the empirical transition probabilities are in all cases 'closer' to the perfect mobility benchmark rather than the perfect immobility benchmark. This finding implies that the fund market is characterized by a considerable degree of mobility. Moreover, in accordance with our initial tests that indicate that persistence vanishes during the 2000s, we find that the data support an increasing similarity of the estimated transition probabilities to the perfect mobility benchmark suggesting that the degree of mobility has increased overtime.

From a methodological point of view the present study provides a formal framework based on a Markovian setting within which persistence can be assessed. Essentially, we explicitly outlined which properties the transition matrices must satisfy in order to be compatible with perfect mobility. Additionally, the employed framework can be seen as a generalization of the widely used "winner-loser" contingency table approach. A distinct advantage of the Markovian analysis is that provides the means for appraising the magnitude of immobility as well as its time evolution. A battery of standard mobility indices act as the vehicle for executing the proposed herein framework.

Future research could modify the Markov Chain allowing for an absorbing state in order to account for fund exits more formally. Other possible modifications would be to correct for the problem of "look ahead bias" (see ter Horst *et. al*., 2001) and also use alternative performance measures such as a Shrinkage Approach (Huij and Verbeek, 2007). In addition, one could explore the driving forces of fund persistence that eventually determine the degree of inertia and its evolution. Financial theory has put forward several explanations for inertia with a prominent one being managerial ability. Thus a fruitful extension of the present study could be the employment of mobility indicators that decompose mobility in terms of several agent characteristics.

References

Agarwal, V. and Naik, N. Y., (2000). "Multi-period Performance Persistence Analysis of Hedge Funds", *Journal of Financial and Quantitative Analysis*, 35, 327-342.

Anderson, T. and L. Goodman, (1957). "Statistical Inference about Markov Chains", *Annals of Mathematical Statistics*, 28, 12-40.

Bangia, A., Diebold, F., Kronimus, A., Schagen, C. and T. Schuermann, (2002). "Ratings Migration and the Business Cycle, with Application to Credit Portfolio Stress Testing", *Journal of Banking and Finance*, 26, 445-474.

Berk, J. B., and R. C. Green (2004), "Mutual fund flows and performance in rational markets," *Journal of Political Economy*, 112: 1269-1295.

Bigard, A., Guillotin, Y. and C. Lucifora, (1998). "Earnings Mobility: An International Comparison of Italy and France", *Review of Income and Wealth*, 44(4), 535-554.

Bollen, N. and J. Busse, (2004). "Short-Term Persistence in Mutual Fund Performance", *Review of Financial Studies*, 18(2), 569-597.

Brown, S. and W. Goetzmann, (1995). "Performance Persistence", *Journal of Finance*, 50(2), 679-698.

Brown, S., Goetzmann, W. and Ibbotson, R., (1999). "Offshore Hedge Funds: Survival and Performance 1989 – 95", *Journal of Business*, 72, 91-117.

Busse, J. and P. Irvine, (2006). "Bayesian Alphas and Mutual Fund Persistence", *Journal of Finance*, 61(5), 2251-2288.

Carhart, M. (1997). "On Persistence in Mutual Fund Performance", *Journal of Finance*, 52(1), 57-82.

Chrinstensen, R. (1990). *Log-Linear Models*, Springer-Verlag, New York.

Cochran, W., (1952). "The χ^2-test of Goodness of Fit", *Annals of Mathematical Statistics*, 23, 315-345.

Elton, E., Gruber, M. and C. Blake, (1996). "The Persistence of Risk-Adjusted Mutual Fund Performance", *Journal of Business*, 69(2), 133-157.

Fama, E. (1965). "The Behavior of Stock Market Prices", *Journal of Business*, 38(1), 34-105.

——————(1970). "Efficient Capital Markets: A Review of Theory and Empirical Work", *Journal of Finance*, 25(2), 383-417.

Fama, E. and K. French, (1993). "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33(1), 3-56.

Fields, G. (2004). "Economic and Social Mobility Really are Multifaceted", Working Paper, *School of Industrial and Labor Relations, Cornell University*.

Geweke, J., Marshall, R. and G. Zarkin, (1986). "Mobility Indices in Continuous Time Markov Chains", *Econometrica*, 54(6), 1407-1423.

Goetzmann, W. and R. Ibbotson, (1994). "Do Winners Repeat? Patterns in Mutual Fund Performance", *Journal of Portfolio Management*, 20, 9-18.

Grinblatt, M. and S. Titman, (1992). "The Persistence of Mutual Funds", *Journal of Finance*, 47, 1977-1984.

Gruber, M. (1996). "Another Puzzle: The Growth in Actively Managed Mutual Funds", *Journal of Finance*, 51, 783-810.

Hendricks, D., Patel, J. and R. Zeckhauser, (1993). "Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974-1988", *Journal of Finance*, 48(1), 93-130.

Huij, J. and M. Verbeek, (2007). "Cross-Sectional Learning and Short-Run Persistence in Mutual Fund Performance", Journal of Banking and Finance, 31, 973-997.

Israel, R., Rosenthal, J. and J. Wei, (2000). "Finding Generators for Markov Chains via Empirical Transition Matrices, with Application to Credit Ratings", *Mathematical Finance*, 11, 245-265.

Jafry, Y. and T. Schuermann, (2004). "Measurement, Estimation and Comparison of Credit Migration Matrices", *Journal of Banking and Finance*, 28, 2603-2639.

Jegadeesh, N. and S. Titman, (1993). "Returns to Buying Winners and Selling Losers: Implication for Stock Market Efficiency", *Journal of Finance*, 48, 65–91.

Jensen, M., (1969). "The Pricing of Capital Assets and Evaluation of Investment Portfolios, *Journal of Business*, 42(2), 167.247.

Malkiel, B., (1995). "Returns from Investing in Equity Mutual Funds 1971-1991", *Journal of Finance*, 50, 549-573.

McQuenn, G. and S. Thorley, (1991). "Are Stock Returns Predictable? A Test Using Markov Chains", *Journal of Finance*, 46(1), 239-263.

Prais, S. (1955). "Measuring Social Mobility", *Journal of the Royal Statistical Society*, Series A, Part I, 118, 56-66.

Shorrocks, A., (1978). "The Measurement of Mobility", *Econometrica*, 46, 1013-1024.

Sommers, P. and J. Conlisk, (1979). "Eigenvalue Immobility Measures for Markov Chains", *Journal of Mathematical Sociology*, 6, 253-276.

Strang, G., (1988). *Linear Algebra and its Applications*, Thomson Publishers.

ter Horst, J., T. Nijman, and M. Verbeek, (2001). "Eliminating Look-Ahead Bias in Evaluating Persistence in Mutual Fund Performance", Journal of Empirical Finance, 8, 345-373.

Theil, H., (1972). *Statistical Decomposition Analysis*, North Holland, Amsterdam.

Trueck, S. and S. Rachev, (2006). "Changes in Migration Matrices and Credit VaR – a New Class of Difference Indices", working paper.

Trueck, S., (2004). "Measures for Comparing Transition Matrices from a Value-at-Risk Perspective", working paper.

Appendix A

A1 Summary mobility and jump indicators

Let p_{ij} denote the transition probability that fund moves from state i in period t to state j in period $t+1$ where n stands for the number of states.

We define:

$R = \left(\sum_{i=j} p_{ij} / \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \right) \cdot 100$, essentially corresponding to the sum of the main diagonal elements, and measures the percentage of funds that have remained at the same relative performance ranking.

$MU = \left(\sum_{i<j} p_{ij} / \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \right) \cdot 100$, that adds-up all elements located in the segment above the main diagonal, and measures the percentage of funds that have *moved up*, that is the percentage of funds that have improved their position in the post-ranking period.

$R = \left(\sum_{i>j} p_{ij} / \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \right) \cdot 100$, that sums all elements located in the segment below the main diagonal, measuring the percentage of funds whose relative position has deteriorated during the evaluation period.

A2 Eigenvalues based mobility indices

Let **P** denote the observed transition matrix and define $tr(\mathbf{P})$, $\det(\mathbf{P})$, and λ_i, $i = 1,2,...,n$ as its trace, determinant and the i-th eigenvalue (arranged in descending order). We employ the following mobility indices: (Prais, 1955; Shorrocks, 1978; Sommers and Conlisk, 1979):

$$M_{PS} = \frac{1}{n-1}(n - tr(\mathbf{P})), \qquad M_{PS} \in [0,1]$$

$$M_S = 1 - |\det(\mathbf{P})| = 1 - \left| \prod_{i=1}^{n} \lambda_i \right|, \quad M_S \in [0,1]$$

$$M_{SC} = 1 - |\lambda_2|, \qquad M_{SC} \in [0,1]$$

$$h = \frac{-\log(2)}{\log|\lambda_2|}$$

A3 Euclidean and singular-vales-based mobility indices

The comparison of any two matrices $(\mathbf{W}, \mathbf{Q})$ with typical elements (w_{ij}, q_{ij}) can be performed using Euclidean cell-by-cell distance measures such as D_{L^1} (Israel *et al.*, 2001), D_{L^2} (Banglia *et al.*, 2002) and $D_{L^{max}}$ (Trueck, 2004) that take the following form:

$$D_{L^1}(\mathbf{W}, \mathbf{Q}) = \sum_{i=1}^{n} \sum_{j=1}^{n} |w_{ij} - q_{ij}| \qquad \in [0,2n]$$

$$\boldsymbol{D_{L^2}}(\mathbf{W}, \mathbf{Q}) = \sum_{i=1}^{n} \sum_{j=1}^{n} (w_{ij} - q_{ij})^2 \qquad \in [0,2n]$$

$$\boldsymbol{D_{L^{max}}}(\mathbf{W}, \mathbf{Q}) = \max_{i,j} |w_{ij} - q_{ij}| \qquad \in [0,1]$$

According to Jafry and Schuermann (2004), the so-called 'mobility' matrix takes the form of $\mathbf{P}_m = \mathbf{P} - \mathbf{I}_n$. Then generalizing the notion of 'amplifying power', initially proposed by Strang (1988), Jafry and Schuermann (2004) advocate the use of the average of all singular values of $\mathbf{P}$, as being more representative of its overall characteristics. In particular they put forward the use of the average of all singular values in the following form:

$$M_{SVD} = \frac{1}{n} \sum_{i=0}^{n} \sqrt{\lambda_i(\mathbf{P}_m' \mathbf{P}_m)}$$

Trueck (2004) and Trueck and Rachev (2006) show that the same principle can be applied to measure the difference between any two (transition) matrices (W, Q) by calculating:

$$D_{SVD}(\mathbf{W}, \mathbf{Q}) = \frac{1}{n} \sum_{i=0}^{n} \sqrt{\lambda_{W,i}(\mathbf{W}'\mathbf{W})} - \frac{1}{n} \sum_{i=0}^{n} \sqrt{\lambda_{Q,i}(\mathbf{Q}'\mathbf{Q})}$$

Appendix B

B.1 Estimating Transition Probabilities

Consider a Markov chain model of a random sequence $\{d_i\}$ with state–space $\mathscr{S} = \{0, 1, \dots, n\}$. In the context of our analysis there are four states ($n = 4$): Funds that belong to the top 25% of performers (best performers), funds that belong to the (26%, 50%) of performers, funds that belong to the (51%, 75%) of performers, and funds that belong to the (76%, 100%) of performers (worst performers). The parameters of the Markov model are its transition probability matrix $\mathbf{P}$ and its initial distribution π_{0i}, $i \in \mathscr{S}$. An obvious solution for estimating the transition probabilities p_{ij} is to resort to the frequency of transitions from i to j in the data. We formalize and illustrate this estimate in this section. We will justify the estimate in terms of statistical theory below after introducing the idea of a maximum likelihood estimate.

Let $N_{ij} \equiv \sum_{l=1}^{n} 1(d_{l-1} = i, d_l = j)$ be the count of all transitions from state i to state j in the data, where $1(a)$ is an indicator function that takes the value 1 when the argument a is true. Let us also define $N_{i\bullet} \equiv \sum_{j \in \mathscr{S}} N_{ij} = \sum_{j=1}^{n} N_{ij}$ as the number of transitions starting from state i in data. Then we will estimate the transition probability p_{ij} by:

$$\hat{p}_{ij} = \frac{N_{ij}}{N_{i\bullet}}$$

Let us briefly explain why this is so. The likelihood function, takes the form

$$\mathbb{L} = p_0 \prod_{i \in \mathcal{S}} \prod_{j \in \mathcal{S}} p_{ij}^{N_{ij}} = p_0 \prod_{i \in \mathcal{S}} \mathbb{L}_i(p_{ij})$$

where $\mathbb{L}_i = \prod_{j \in \mathcal{S}} p_{ij}^{N_{ij}}$ depends on the elements of the i-th row of the transition matrix. The log-likelihood then becomes:

$$\ln \mathbb{L} = \mathcal{L} = \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} N_{ij} \ln p_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} N_{ij} \ln p_{ij} \quad \text{s.t.} \quad \sum_{j \in \mathcal{S}} p_{ij} = \sum_{j=1}^{n} p_{ij} = 1, \ p_{ij} \geq 0$$

The FOC for likelihood maximization yield:

$$\hat{p}_{ij} = \frac{N_{ij}}{\sum_{j \in \mathcal{S}} N_{ij}} = \frac{N_{ij}}{\sum_{j=1}^{n} N_{ij}} = \frac{N_{ij}}{N_{i\bullet}},$$

as the as the asymptotically unbiased and normally distributed ML estimator of p_{ij} (Anderson and Goodman, 1957). One can think of this as multinomial likelihoods from random samples.

B.2 Tests for a Specified Transition Matrix

One may be interested in testing whether or not the estimated transition matrix is equal to an exogenously given transition matrix, *i.e.*, whether or not $p_{ij} = p_{ij}^0$ holds for all $i, j \in \mathcal{S}$. The appropriate test statistic, known as χ^2 test of goodness of fit (Cochran 1952; Anderson and Goodman, 1957), reads

$$\mathcal{Q}^{(U)} = \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} N_{i\bullet} \frac{\left(\hat{p}_{ij} - p_{ij}^0\right)^2}{p_{ij}^0} \sim \chi^2\left(n(n-1)\right),$$

where again n is the number of states in the state-space considered. Alternatively, one could employ a likelihood ratio test, given by:

$$\mathcal{LR}^{(U)} = \sum_{i \in \mathcal{S}} \sum_{i \in \mathcal{S}} N_{ij} \ln\left(\frac{\hat{p}_{ij}}{p_{ij}^0}\right) \sim \chi^2\left(n(n-1)\right).$$

B.3 Tests of Time-Homogeneity

A test of time homogeneity (time stationarity) is appropriate for deciding whether the transition probabilities of the first-order Markov chain can be assumed constant over time. This test can be performed by dividing the entire sample period comprising T transitions into M mutually exclusive and exhaustive sub-periods ($t = 1, 2, ..., M$; $M \leq T$) and comparing the transition matrices estimated from each of the M sub-samples to the matrix estimated from the entire sample. More specifically, one can test:

$$\mathcal{H}_0^{(T)} : p_{ij}(t) = p_{ij} \ \forall t \ (t = 1, 2, ..., M)$$

Against the alternative of transition probabilities differing between periods:

$$\mathcal{H}_1^{(T)} : p_{ij}(t) \neq p_{ij} \text{ for some } t,$$

where p_{ij} is the probability of transition from state i to state j for the full sample ("pooled" across M periods), and $p_{ij}(t)$ the corresponding transition probability from the t-th sub-sample. Assuming that there are at least two non-zero transition probabilities in each row (i) of the transition matrix for the entire sample, one can either perform a likelihood ratio test of the form:[13]

$$\mathcal{LR}^{(T)} = \sum_{t=1}^{M} \sum_{i \in \mathcal{S}} \sum_{i \in \mathcal{S}} N_{ij}(t) \ln \left(\frac{\hat{p}_{ij}(t)}{\hat{p}_{ij}} \right) \sim \chi^2 \left(n(n-1)M \right),$$

where n denotes the number of elements in $\mathcal{S}$ (the number of states). Similarly, one can employ a Pearson-type χ^2 statistic:

$$\mathcal{Q}^{(T)} = \sum_{t=1}^{M} \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} N_{i\bullet}(t) \frac{\left(\hat{p}_{ij}(t) - \hat{p}_{ij} \right)}{\hat{p}_{ij}} \sim \chi^2 \left(n(n-1)M \right),$$

The above test statistics have an asymptotic χ^2 distribution with degrees of freedom equal to the number of additional independent restrictions imposed under $\mathcal{H}_0^{(T)}$ relative to $\mathcal{H}_1^{(T)}$.

---

[13] The transition probabilities $\hat{p}_{ij}$ are assumed to be mutually independent across sub-samples under $\mathcal{H}_0^{(T)}$, hence the n^2 parameters can be estimated similar to the formula in text $\hat{p}_{ij}(t) = N_{ij}(t) / N_{i\bullet}(t)$.

Tables

Table B.1: Averages for Mutual Fund Returns and Risk Factors by Year

Year	R	RMRF	SMB	HML	MOM
1982	1.270	0.862	0.605	0.813	**2.305**
1983	0.890	1.054	0.927	1.280	**-0.629**
1984	-0.820	-0.449	-0.662	1.544	**0.764**
1985	1.620	1.740	0.025	0.02	**0.904**
1986	0.730	0.824	-0.762	0.649	**0.705**
1987	0.010	0.096	-0.711	-0.259	**0.006**
1988	0.590	0.877	0.415	1.001	**-0.430**
1989	1.300	1.480	-0.841	-0.347	**1.966**
1990	-0.910	-1.023	-1.248	-0.981	**1.677**
1991	2.210	2.075	1.021	-0.900	**0.818**
1992	0.460	0.457	0.574	1.695	**0.229**
1993	0.680	0.695	0.465	1.377	**1.727**
1994	-0.470	-0.343	-0.117	-0.068	**0.297**
1995	1.780	2.130	-0.465	0.076	**1.165**
1996	1.070	1.232	-0.120	0.165	**0.460**
1997	1.450	1.888	-0.318	0.743	**0.810**
1998	1.330	1.507	-1.768	-0.792	**1.734**
1999	1.700	1.590	1.112	-2.415	**2.291**
2000	-0.860	-1.311	-3.660	5.095	**1.640**
2001	-1.470	-1.128	1.645	1.273	**-0.343**
2002	-2.260	-1.914	0.356	1.018	**2.375**
2003	2.110	2.376	1.684	0.267	**-1.520**
2004	0.720	0.953	0.420	0.681	**0.028**
2005	0.320	0.375	-0.128	0.679	**1.174**
Overall	0.350	0.668	-0.064	0.525	**0.840**

Note: all returns are expressed as percentages.

Table B.2. Descriptive Statistics for Entry and Exit Rates (1982 – 2005)

Statistic	Entry Rate	Exit Rate
Mean	1.26	0.24
Median	1.15	0.13
Maximum	4.33	4.32
Minimum	0.00	0.00
Standard Deviation	0.90	0.39
Skeweness	0.87	4.72
Kurtosis	3.69	43.51

Note: The table reports descriptive statistics of the annual entry and exit rates (in percent) over the whole sample period 1982 to 2005.

TableB.3: Stationarity Tests

Stationarity Test: $H_0 : p_{ij}^{(t)} = p_{ij}, \; \forall t$		
LR Test	Pearson-type Test	d.f.
Panel A: Full Sample Results		
Panel A.1: Time Homogeneity Tests (two sub-periods)		
49.872	48.620	12
[0.000]	[0.000]	
Panel A.2: Time Homogeneity Tests (seven sub-periods)		
152.086	149.861	72
[0.000]	[0.000]	
Panel B: Stationarity Tests: 1985-1993, 1994-2005		
Panel B.1: Time Homogeneity Tests 1985-1993 (three sub-periods)		
38.642	38.438	24
[0.030]	[0.031]	
Panel B.2: Time Homogeneity Tests 1994-2005 (three sub-periods)		
58.736	58.944	24
[0.000]	[0.000]	

Notes: Panel A reports tests for the assumption of stationarity of the transition probabilities comparing the estimates of transition probabilities from two sub-periods (Panel A.1 the sub-periods being 1985-1996, 1997-2005) and seven sub-periods (Panel A.2) with the transition matrix estimated from the full sample. Panel B, reports similar tests for two sub-periods: 1985-1993 and 1994-2005. Panel B.1 reports the same test of time homogeneity of the transition probabilities, comparing the transition matrix estimated over 1985-1993 with those estimated in three different sub-periods. Panel B.2 reports similar tests for the period 1994-2005. Each test statistic is asymptotically distributed as χ^2 with degrees of freedom given in the last column of the table. *P*-values are given in square brackets.