Financial Econometrics
Shanghai Jiao Tong University
Summer School 2026
Empirical Exercise 14.2
With Solutions

Fatih Kansoy

Reference: Stock, J.H. and Watson, M.W., Introduction to Econometrics (4th Edition, Global Edition)

Data: `Stock_Returns_1931_2002.csv` *with columns* `time` *(year),* `Month`*,* `ExReturn` *(monthly percentage excess return), and* `ln_DivYield` *(100× log dividend yield). Use 1932:M1–2002:M12.*

Empirical Exercise 14.2. Read the boxes "Can You Beat the Market? Part I" and "Can You Beat the Market? Part II" in Chapter 14. These boxes study whether monthly excess returns on a broad stock index can be forecast from past returns and from the dividend yield, and relate the answer to the efficient markets hypothesis. The dataset `Stock_Returns_1931_2002.csv` is an extended version of the data described in those boxes. It contains the columns `time` (year), `Month`, `ExReturn` (the monthly excess return on the CRSP value-weighted index, measured in percentage points per month), and `ln_DivYield` (100 times the logarithm of the dividend yield). In this exercise you will re-estimate the boxes' regressions over the longer sample 1932:M1–2002:M12.

(a) Repeat the calculations reported in Table 14.2, using regressions estimated over the 1932:M1–2002:M12 sample period. That is, estimate AR(1), AR(2) and AR(4) models for the excess return `ExReturn` and report the estimated coefficients, the intercept, the F-statistic (with its p-value) testing that all lag coefficients are zero, and R^2.

(b) Repeat the calculations reported in Table 14.6, using regressions estimated over the 1932:M1–2002:M12 sample period. That is, estimate autoregressive distributed lag

models of `ExReturn` that use the dividend yield `ln_DivYield` (in first differences and in levels, following the specifications in the box) as an additional predictor.

(c) Is the variable ln(*dividend yield*), that is `ln_DivYield`, highly persistent? Explain how you reached your conclusion.

(d) Construct pseudo out-of-sample forecasts of excess returns over the 1983:M1–2002:M12 period, using forecasting regressions that begin in 1932:M1. Compare the root mean squared forecast error (RMSFE) of the candidate models with that of a "zero forecast" (always forecasting a zero excess return) and a "constant forecast" (a recursively estimated model that contains only an intercept).

(e) Do the results in (a) through (d) suggest any important changes to the conclusions reached in the boxes? Explain.

Solution:

Data Overview

The dataset `Stock_Returns_1931_2002.csv` contains monthly observations on excess returns and the dividend yield for the CRSP value-weighted stock index. The estimation sample used throughout is 1932:M1–2002:M12, with earlier observations reserved for constructing lagged variables.

Variable	Description	Units
`time`	Year of the observation	calendar year
`Month`	Month of the observation	1–12
`ExReturn`	Monthly excess return on the CRSP index	% per month
`ln_DivYield`	$100 \times \ln(\textit{dividend yield})$	index points

Note: The excess return is the return from buying the index at the end of the previous month and selling it at the end of the current month, less the return on a safe Treasury bill over the same month. The boxes in the textbook study the 1960:M1–2002:M12 sample. Here we use the longer 1932:M1–2002:M12 sample, so the coefficients differ slightly from the textbook tables but the economic message is the same.

Key Notation and Formulas

An autoregression of order p for the excess return R_t is:

$$R_t = \beta_0 + \beta_1 R_{t-1} + \beta_2 R_{t-2} + \cdots + \beta_p R_{t-p} + u_t.$$

The F-statistic tests the joint null hypothesis that all autoregressive coefficients are zero:

$$H_0 : \ \beta_1 = \beta_2 = \cdots = \beta_p = 0.$$

The root mean squared forecast error over a pseudo out-of-sample period of P forecasts is:

$$\text{RMSFE} = \sqrt{\frac{1}{P}\sum_{t=1}^{P}\left(R_t - \widehat{R}_t\right)^2},$$

where $\widehat{R}_t$ is the forecast computed using only information available before period t.

(a) Autoregressive Models of Excess Returns (Table 14.2 Extended)

Sample period: 1932:M1–2002:M12.

We estimate AR(1), AR(2) and AR(4) models for `ExReturn`. The results are:

Regressor	**(1) AR(1)**	**(2) AR(2)**	**(3) AR(4)**
Excess Return$_{t-1}$	0.098	0.102	0.099
	(0.061)	(0.061)	(0.058)
Excess Return$_{t-2}$		0.040	0.029
		(0.057)	(0.054)
Excess Return$_{t-3}$			0.098
			(0.054)
Excess Return$_{t-4}$			0.006
			(0.046)
Intercept	0.524	0.543	0.590
	(0.181)	(0.186)	(0.199)
F-statistic (all lags)	2.61	1.51	1.41
(p-value)	(0.11)	(0.22)	(0.23)
R^2	0.009	0.009	0.016

Note: standard errors in parentheses.

Interpretation:

- In every specification the coefficients on lagged returns are small and statistically insignificant, with t-statistics below 2 in absolute value.
- The F-statistics testing that all lag coefficients are zero do not reject the null at the 5% level in any model (p-values of 0.11, 0.22 and 0.23).
- The R^2 values are tiny (under 2%), so past returns explain almost none of the variation in this month's excess return.
- The point estimate on the first lag (about 0.10) is slightly larger than in the textbook's 1960–2002 sample, but it remains insignificant once the standard error is taken into account.

These results echo the box "Can You Beat the Market? Part I": past returns provide essentially no useful information for forecasting future excess returns, consistent with the efficient markets hypothesis in its weak form.

(b) Distributed Lag Models with the Dividend Yield (Table 14.6 Extended)

We add the (log) dividend yield `ln_DivYield` to the autoregression, following the specifications used in the box "Can You Beat the Market? Part II". As in the textbook, the dividend yield enters in first differences (columns 1 and 2), because it is highly persistent, and once in levels (the ADL(1,1) specification) to match the economic reasoning that links returns to the level of the yield.

Specification	**Finding**
ADL(1,1) with Δ`ln_DivYield`	lag coefficients insignificant
ADL(2,2) with Δ`ln_DivYield`	lag coefficients insignificant
ADL(1,1) with level of `ln_DivYield`	coefficient borderline; inference suspect

Interpretation:

- When the dividend yield enters in *first differences*, the individual t-statistics and the joint F-statistics fail to reject the null hypothesis of no predictability, just as in the textbook box. The added predictor does not help.
- When the dividend yield enters in *levels*, the coefficient can appear borderline significant against the usual 1.96 critical value. However, because `ln_DivYield` is highly persistent (see part (c)), the usual normal distribution for the t-statistic is not reliable, so the 1.96 critical value may be inappropriate and the apparent significance should be treated with caution.
- The R^2 values remain very small in every specification, confirming that the dividend yield adds little forecasting power over the longer sample.

The longer 1932–2002 sample therefore reproduces the box's central caution: any in-sample "predictability" from the level of the dividend yield is fragile and is undermined by the persistence of the regressor.

(c) Is the Log Dividend Yield Highly Persistent?

Yes. The variable `ln_DivYield` is highly persistent.

How to reach this conclusion:

1. **Autocorrelation.** The first-order autocorrelation of `ln_DivYield` is very close to 1. The series moves slowly and smoothly, so its current value is almost equal to its value last month.

2. **Unit root test.** A Dickey–Fuller test (with an intercept) on `ln_DivYield` fails to reject the null hypothesis of a unit root, even at the 10% significance level. The estimated largest autoregressive root is essentially 1.

3. **Practical consequence.** Because the regressor is so persistent, the distribution of the t-statistic on its level is non-standard. This is exactly why the box recommends entering the dividend yield in first differences, and why the apparent significance of the level coefficient in part (b) must be interpreted cautiously.

Key point: A failure to reject a unit root does not prove the series has one, but it does confirm that `ln_DivYield` is highly persistent. Standard inference based on the normal distribution is therefore unreliable when the level of this variable is used as a regressor.

(d) Pseudo Out-of-Sample Forecasts, 1983:M1–2002:M12

We construct pseudo out-of-sample forecasts of excess returns over 1983:M1–2002:M12. Each forecasting model is estimated recursively, using only data from 1932:M1 up to the month before the forecast, and is then used to forecast one month ahead. We compare the AR and constant-only models with the two simple benchmarks described in the box.

[See Figure: Pseudo out-of-sample forecasts of excess returns, 1983:M1–2002:M12]

Model	**RMSFE**
Zero forecast	4.28
Constant forecast	**4.26**
AR(1)	4.28
AR(2)	4.27
AR(4)	4.29

Interpretation:

- The most accurate forecast comes from the **constant-term forecast** (RMSFE = 4.26), that is, a recursively estimated model that contains only an intercept.
- The "zero forecast" (always forecasting an excess return of zero) and the AR models are all essentially as good, or slightly worse: their RMSFEs cluster between 4.27 and 4.29.
- The differences across models are very small, well within sampling noise. Adding lags of returns does not improve out-of-sample accuracy. This is consistent with the statistical insignificance of the AR coefficients found in part (a).

Key finding: The simplest models (a constant forecast, or simply forecasting zero) do at least as well out of sample as the autoregressive models. Past returns provide no usable forecasting advantage.

(e) Do the Results Change the Conclusions of the Boxes?

No. The results over the longer 1932:M1–2002:M12 sample do not overturn the conclusions reached in the boxes.

1. **Past returns do not forecast future returns.** The AR(1), AR(2) and AR(4) models in part (a) have insignificant lag coefficients, statistically insignificant F-statistics, and negligible R^2. This matches Part I of the box.

2. **The dividend yield adds little.** In part (b), when the dividend yield enters in first differences it is insignificant. When it enters in levels, any apparent significance is unreliable because the regressor is highly persistent (part (c)). This matches Part II of the box.

3. **Out-of-sample, simple is best.** In part (d), the constant forecast (and even the zero forecast) is as accurate as the autoregressive models. The extra predictors buy no out-of-sample gain.

Bottom line: Over the longer 1932–2002 sample, excess returns remain essentially unpredictable using past returns or the dividend yield. The evidence continues to support the efficient markets hypothesis, and the conclusions of the two boxes are unchanged.

Summary: Key Takeaways

Question	Finding	Interpretation
AR models of returns	Lags insignificant	Past returns do not forecast
Dividend yield (differences)	Insignificant	No added predictability
Dividend yield (level)	Borderline, suspect	Persistence invalidates inference
Is `ln_DivYield` persistent?	Yes (near unit root)	Standard t-test unreliable
Best out-of-sample model	Constant forecast (4.26)	Simple beats complex
Change conclusions?	**No**	Markets look efficient

The Big Picture

This exercise illustrates several important ideas in time series forecasting:

1. **Efficient markets and unpredictability:** If excess returns were easily forecast, traders would act on that information and drive prices to the point where the expected excess return is zero. The near-zero predictability we find is exactly what this reasoning leads us to expect.

2. **Persistent regressors:** When a predictor such as the log dividend yield is highly persistent (close to a unit root), the standard normal critical value of 1.96 for its t-statistic can be misleading. Apparent significance may be a statistical artefact, which is why the dividend yield is entered in first differences.

3. **In-sample fit versus out-of-sample accuracy:** A model can show a slightly higher R^2 in sample yet forecast no better out of sample. The pseudo out-of-sample exercise in part (d), where a constant forecast wins, is a clean reminder that out-of-sample accuracy is the honest test of a forecasting model.

4. **Robustness across samples:** Extending the sample from 1960–2002 to 1932–2002 changes the point estimates a little but not the overall conclusion. Findings that survive a longer sample are more credible than those that depend on one particular period.

Calculations performed using Stata over the 1932:M1–2002:M12 sample.