Exchange Rate Models and Spurious Regressions

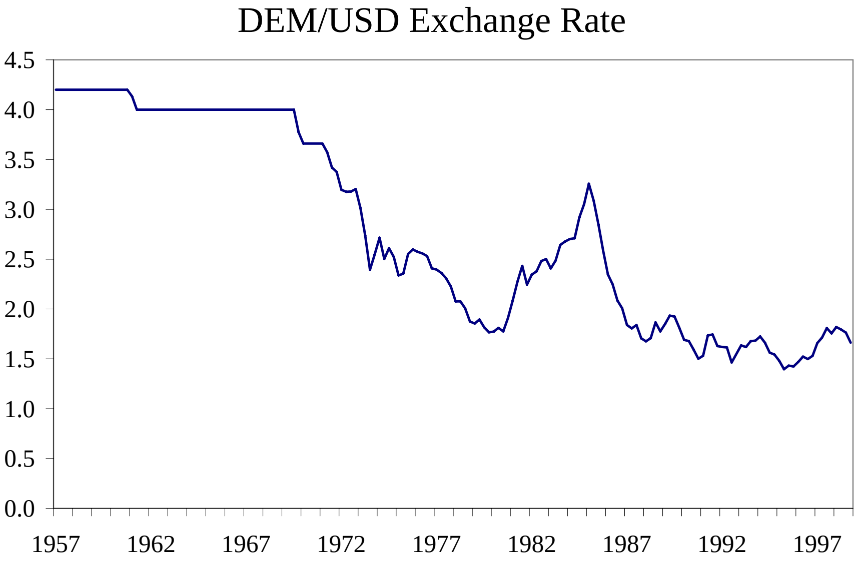
Jón Steinsson and Emi Nakamura

UC Berkeley

April 2019

"One of the most remarkable facts about G3 exchange rates is that they are so seemingly immune to systematic empirical explanation."

– Kenneth Rogoff


DEM/USD Exchange Rate
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1957
1962
1967
1972
1977
1982
1987
1992
1997

1970's Monetary Model of Exchange Rates

- Purchasing Power Parity

$$e_t = p_t - p_t^*$$

1970's Monetary Model of Exchange Rates

- Purchasing Power Parity

$$e_t = p_t - p_t^*$$

- Money demand

$$m_t - p_t = \phi_y y_t - \phi_i i_t$$

$$m_t^* - p_t^* = \phi_y y_t^* - \phi_i i_t^*$$

1970's Monetary Model of Exchange Rates

- Purchasing Power Parity

$$e_t = p_t - p_t^*$$

- Money demand

$$m_t - p_t = \phi_y y_t - \phi_i i_t$$

$$m_t^* - p_t^* = \phi_y y_t^* - \phi_i i_t^*$$

- Combining money demand

$$p_t - p_t^* = (m_t - m_t^*) - \phi_y (y_t - y_t^*) + \phi_i (i_t - i_t^*)$$

1970's Monetary Model of Exchange Rates

- Purchasing Power Parity

$$e_t = p_t - p_t^*$$

- Money demand

$$m_t - p_t = \phi_y y_t - \phi_i i_t$$

$$m_t^* - p_t^* = \phi_y y_t^* - \phi_i i_t^*$$

- Combining money demand

$$p_t - p_t^* = (m_t - m_t^*) - \phi_y(y_t - y_t^*) + \phi_i(i_t - i_t^*)$$

- Exchange rate and "fundamentals":

$$e_t = (m_t - m_t^*) - \phi_y(y_t - y_t^*) + \phi_i(i_t - i_t^*)$$

Frenkel (1976)

- Sample: German Mark, February 1920 - November 1923.
- Hyperinflation: Ignore a bunch of terms.

$$e_t = (m_t - m_t^*) - \phi_y(y_t - y_t^*) + \phi_i(i_t - i_t^*)$$

$$e_t = m_t - \phi_i(i_t - i_t^*)$$

$$\log S = -5.135 + 0.975 \log M + 0.591 \log \pi$$
$$(0.731) \quad (0.050) \qquad\qquad (0.073)$$

$$R^2 = 0.994; \text{ s.e.} = 0.241; \text{ D.W.} = 1.91.$$

Source: Frenkel (1976).

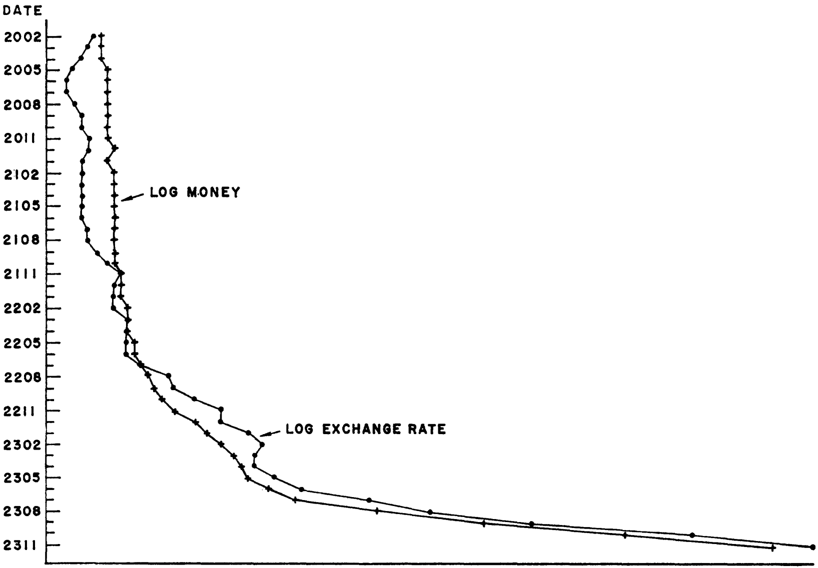


Fig. 1.

Source: Frenkel (1976).

Frankel (1979)

- Sample: DEM/USD, July 1974 - February 1978.

$$e_t = \phi_0 + \phi_m(m_t - m_t^*) - \phi_y(y_t - y_t^*)$$
$$+\phi_i(i_t - i_t^*) + \phi_\pi(\pi_t^e - \pi_t^{e*}) + \epsilon_t$$

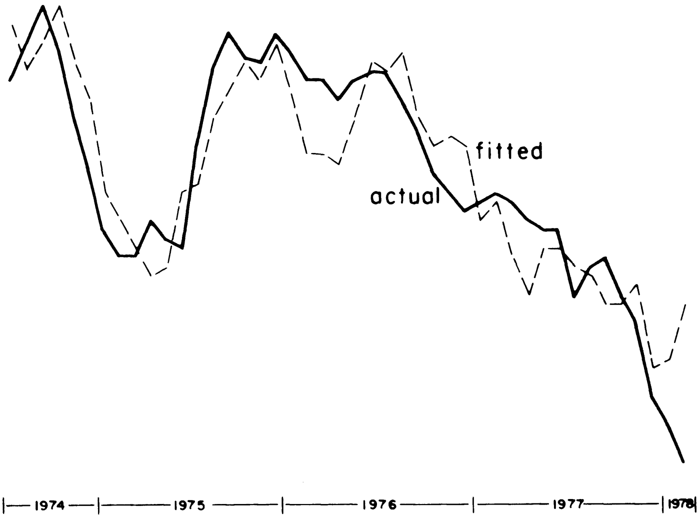


FIGURE 1. PLOT OF (*log* OF) MARK/DOLLAR RATE, *OLS* REGRESSION FROM TABLE 1

Source: Frankel (1979).

TABLE 1—TEST OF REAL INTEREST DIFFERENTIAL HYPOTHESIS
(Sample: July 1974–February 1978)

Technique	Constant	$m - m_1^*$	$y - y^*$	$r - r^*$	$\pi - \pi^*$	R^2	*D.W.*	$\hat{\rho}$	Number of Observations
OLS	1.33	.87	−.72	−1.55	28.65	.80	.76		44
	(.10)	(.17)	(.22)	(1.94)	(2.70)				
CORC	.80	.31	−.33	−.259	7.72	.91		.98	43
	(.19)	(.25)	(.20)	(1.96)	(4.47)				
INST	1.39	.96	−.54	−4.75	27.42		1.00		42
	(.08)	(.14)	(.18)	(1.69)	(2.26)				
FAIR	1.39	.97	−.52	−5.40	29.40			.46	41
	(.12)	(.21)	(.22)	(2.04)	(3.33)				

Note: Standard errors are shown in parentheses.
Definitions: Dependent Variable (*log* of) Mark/Dollar Rate.

CORC = Iterated Cochrane-Orcutt.

INST = Instrumental variables for expected inflation differential are Consumer Price Index (*CPI*) inflation differential (average for past year), industrial Wholesale Price Index (*WPI*) inflation differential (average for past year), and long-term commercial bond rate differential.

FAIR = Instrumental variables are industrial *WPI* inflation differential and lagged values of the following: exchange rate, relative industrial production, short-term interest differential, and expected inflation differential. The method of including among the instruments lagged values of all endogenous and included exogenous variables, in order to insure consistency while correcting for first-order serial correlation, is attributed to Ray Fair.

$m - m^*$ = *log* of German M_1/*U.S.* M_1

$y - y^*$ = *log* of German production/*U.S.* production

$r - r^*$ = Short–term German-*U.S.* interest differential

$(r - r^*)_{-1}$ = Short-term German-*U.S.* interest differential lagged

$\pi - \pi^*$ = Expected German-*U.S.* inflation differential, proxied by long-term government bond differential.

Source: Frankel (1979).

Meese and Rogoff (1983)

- Do the monetary models of exchange rates fit out of sample?

Meese and Rogoff (1983)

- Do the monetary models of exchange rates fit out of sample?
- Generalized monetary model:

$$e_t = \phi_0 + \phi_m(m_t - m_t^*) + \phi_y(y_t - y_t^*) + \phi_i(i_t - i_t^*)$$
$$+\phi_\pi(\pi_t^e - \pi_t^{e*}) + \phi_{TB} TB_t + \phi_{TB*} TB_t^* + \epsilon_t$$

Meese and Rogoff (1983)

- Auto-regressive model

$$e_t = \phi_0 + \sum_{j=1}^{J} \phi_j e_{t-j} + \epsilon_t$$

Meese and Rogoff (1983)

- Auto-regressive model

$$e_t = \phi_0 + \sum_{j=1}^{J} \phi_j e_{t-j} + \epsilon_t$$

- Vector auto-regressive model

$$e_t = \phi_0 + \sum_{j=1}^{J} \phi_j e_{t-j} + \sum_{j=1}^{J} \Phi_j X_{t-j} + \epsilon_t$$

Meese and Rogoff (1983)

- Auto-regressive model

$$e_t = \phi_0 + \sum_{j=1}^{J} \phi_j e_{t-j} + \epsilon_t$$

- Vector auto-regressive model

$$e_t = \phi_0 + \sum_{j=1}^{J} \phi_j e_{t-j} + \sum_{j=1}^{J} \Phi_j X_{t-j} + \epsilon_t$$

- Random Walk model

$$E_t e_{t+j} = e_t$$

Meese and Rogoff (1983)

- Sample period: March 1973 - June 1981
- Forecasts based on rolling regression starting November 1976
- Forecast horizons: 1, 6 and 12 months

- Sample period: March 1973 - June 1981
- Forecasts based on rolling regression starting November 1976
- Forecast horizons: 1, 6 and 12 months
- Measure of out-of-sample accuracy: RMSE

$$\left\{ \sum_{s=0}^{N_k - 1} [F(t+s+k) - A(t+s+k)]^2 / N_k \right\}^{1/2}$$

Meese and Rogoff (1983)

In structural models:

- Use actual realized future values of explanatory variables (as opposed to also forecasting explanatory variables)

Meese and Rogoff (1983)

In structural models:

- Use actual realized future values of explanatory variables (as opposed to also forecasting explanatory variables)

Two possible stories:

- Hard to predict exchange rate because it is hard to predict variables that it depends on

MEESE AND ROGOFF (1983)

In structural models:

- Use actual realized future values of explanatory variables (as opposed to also forecasting explanatory variables)

Two possible stories:

- Hard to predict exchange rate because it is hard to predict variables that it depends on
- Hard to find any systematic relantionship between exchange rates and other variables

Table 1

Root mean square forecast errors.[a]

Exchange rate	Model: / Horizon	Random walk	Forward rate	Univariate autoregression	Vector autoregression	Frenkel–Bilson[b]	Dornbusch–Frankel[b]	Hooper–Morton[b]
	1 month	3.72	3.20	3.51	5.40	3.17	3.65	3.50
$/mark	6 months	8.71	9.03	12.40	11.83	9.64	12.03	9.95
	12 months	12.98	12.60	22.53	15.06	16.12	18.87	15.69
	1 month	3.68	3.72	4.46	7.76	4.11	4.40	4.20
$/yen	6 months	11.58	11.93	22.04	18.90	13.38	13.94	11.94
	12 months	18.31	18.95	52.18	22.98	18.55	20.41	19.20
	1 month	2.56	2.67	2.79	5.56	2.82	2.90	3.03
$/pound	6 months	6.45	7.23	7.27	12.97	8.90	8.88	9.08
	12 months	9.96	11.62	13.35	21.28	14.62	13.66	14.57
Trade-weighted dollar	1 month	1.99	N.A.	2.72	4.10	2.40	2.50	2.74
	6 months	6.09	N.A.	6.82	8.91	7.07	6.49	7.11
	12 months	8.65	14.24	11.14	10.96	11.40	9.80	10.35

[a]Approximately in percentage terms.

[b]The three structural models are estimated using Fair's instrumental variable technique to correct for first-order serial correlation.

($/Mark 1-month number should be 3.17 not 3.72, see Table 3)

Source: Meese and Rogoff (1983).

Meese and Rogoff (1983)

- Nothing beats random walk out of sample

- Nothing beats random walk out of sample
- Stronger than just lack of predictability

  (since they use realized future values of explanatory variables)
- Nothing even explains exchange rates!!!

Meese and Rogoff (1983)

Rogoff (2001) recounts:

> *For a long time, no one did believe us. The editor of the American Economic Reivew (Robert Clower) sent our manuscript back in return mail with a scathing letter saying that the results are obviously garbage and if we wish to remain in the economics profession, we had better develop a more positive attitude. ... One then young and now preeminent MIT macroeconomist, when told the findings, forcefully commented (with a French accent) "You just cannot possibly have done it right."*

Meese and Rogoff (1983)

Rogoff (2001) recounts:

> *For a long time, no one did believe us. The editor of the American Economic Reivew (Robert Clower) sent our manuscript back in return mail with a scathing letter saying that the results are obviously garbage and if we wish to remain in the economics profession, we had better develop a more positive attitude. ... One then young and now preeminent MIT macroeconomist, when told the findings, forcefully commented (with a French accent) "You just cannot possibly have done it right."*

As of April 2019: 4776 Google scholar citations

Two Lessons

1. Economics lesson:
    - Exchange rate dominated by unpredicatable shocks (unpredictable capital flows?)
    - Exchange rate very forward looking variable

Two Lessons

1. Economics lesson:
   - Exchange rate dominated by unpredicatable shocks (unpredictable capital flows?)
   - Exchange rate very forward looking variable
2. Econometric lesson:
   - Beware regressing very persistent variable on another very persistent variable

Simplified Engel and West (2005)

- Uncovered interest rate parity:

$$i_t = i_t^* + E_t e_{t+1} - e_t$$

- Returns should be equalized across countries
- If interest rate is higher abroad, exchange rate should fall enough on average to equalize returns

  (e_t is domestic currency price of foreign currency)

Simplified Engel and West (2005)

Rearranging and solving forward:

$$i_t = i_t^* + E_t e_{t+1} - e_t$$

$$e_t = (i_t^* - i_t) + E_t e_{t+1}$$

$$e_t = (i_t^* - i_t) + \sum_{j=1}^{\infty} E_t(i_{t+j}^* - i_{t+j}) + \lim_{j\to\infty} E_t e_{t+j}$$

Simplified Engel and West (2005)

What determines the change in the exchange rate:

$$e_{t+1} - e_t = -(i_t^* - i_t) + \sum_{j=1}^{\infty} \Delta E_{t+1}(i_{t+j}^* - i_{t+j}) + \lim_{j \to \infty} \Delta E_{t+1} e_{t+j}$$

where $\Delta E_{t+1} x_{t+j} = E_{t+1} x_{t+j} - E_t x_{t+j}$ (time $t+1$ news about x_{t+j})

Simplified Engel and West (2005)

What determines the change in the exchange rate:

$$e_{t+1} - e_t = -(i_t^* - i_t) + \sum_{j=1}^{\infty} \Delta E_{t+1}(i_{t+j}^* - i_{t+j}) + \lim_{j\to\infty} \Delta E_{t+1} e_{t+j}$$

where $\Delta E_{t+1} x_{t+j} = E_{t+1} x_{t+j} - E_t x_{t+j}$ (time $t+1$ news about x_{t+j})

- Two components:
  - Current interest rate differencial
  - News about all future interest rate differentials

Simplified Engel and West (2005)

What determines the change in the exchange rate:

$$e_{t+1} - e_t = -(i_t^* - i_t) + \sum_{j=1}^{\infty} \Delta E_{t+1}(i_{t+j}^* - i_{t+j}) + \lim_{j \to \infty} \Delta E_{t+1} e_{t+j}$$

where $\Delta E_{t+1} x_{t+j} = E_{t+1} x_{t+j} - E_t x_{t+j}$ (time $t+1$ news about x_{t+j})

- Two components:
  - Current interest rate differencial
  - News about all future interest rate differentials
- Not so implausible that the variance of the latter is huge compared to the former

Simplified Engel and West (2005)

$$e_{t+1} - e_t = -(i_t^* - i_t) + \sum_{j=1}^{\infty} \Delta E_{t+1}(i_{t+j}^* - i_{t+j}) + \lim_{j \to \infty} \Delta E_{t+1} e_{t+j}$$

- But $(i_t^* - i_t)$ not only thing observed

Simplified Engel and West (2005)

$$e_{t+1} - e_t = -(i_t^* - i_t) + \sum_{j=1}^{\infty} \Delta E_{t+1}(i_{t+j}^* - i_{t+j}) + \lim_{j \to \infty} \Delta E_{t+1} e_{t+j}$$

- But $(i_t^* - i_t)$ not only thing observed
- Movements in longer-term bonds allow one to back out estimates of

$$\sum_{j=1}^{\infty} \Delta E_{t+1}(i_{t+j}^* - i_{t+j})$$

at least up to j = 40 quarters (and assuming EHTS)

Simplified Engel and West (2005)

$$e_{t+1} - e_t = -(i_t^* - i_t) + \sum_{j=1}^{\infty} \Delta E_{t+1}(i_{t+j}^* - i_{t+j}) + \lim_{j\to\infty} \Delta E_{t+1} e_{t+j}$$

- But $(i_t^* - i_t)$ not only thing observed
- Movements in longer-term bonds allow one to back out estimates of

$$\sum_{j=1}^{\infty} \Delta E_{t+1}(i_{t+j}^* - i_{t+j})$$

  at least up to j = 40 quarters (and assuming EHTS)
- $\lim_{j\to\infty} \Delta E_{t+1} e_{t+j}$ still a potential problem
- But in real terms PPP should hold in the very long run (Clarida-Luo 14; Engel 15)

In-sample versus Out-of-sample

- Why was Frankel's in-sample inference so much stronger than Meese-Rogoff's out-of-sample inference?

In-sample versus Out-of-sample

- Why was Frankel's in-sample inference so much stronger than Meese-Rogoff's out-of-sample inference?
- Suggests that something is wrong with in-sample inference (This is a general concern)

Spurious Regressions

- Monetary model of exchange rate:

$$e_t = \phi_0 + \phi_f f_t + \epsilon_t$$

- Both e_t and f_t have a unit-root.
- Granger and Newbold (1974):
  - Usual methods massively understate standard errors

As a preliminary, we looked at the regression

$$Y_t = \beta_0 + \beta_1 X_t,$$

where Y_t and X_t were, in fact, generated as *independent* random walks each of length 50. Table 1 shows values of

$$S = \frac{|\hat{\beta}_1|}{\widehat{S.E.}(\hat{\beta}_1)},$$

the customary statistic for testing the significance of β_1, for 100 simulations.

Table 1
Regressing two independent random walks.

S:	0–1	1–2	2–3	3–4	4–5	5–6	6–7	7–8
Frequency:	13	10	11	13	18	8	8	5
S:	8–9	9–10	10–11	11–12	12–13	13–14	14–15	15–16
Frequency:	3	3	1	5	0	1	0	1

Source: Granger and Newbold (1974).

Table 2

Regressions of a series on m independent 'explanatory' series.

Series either all random walks or all A.R.I.M.A. (0, 1, 1) series, or changes in these. $Y_0 = 100$, $Y_t = Y_{t-1} + a_t$, $Y_t' = Y_t + kb_t$; $X_{j,0} = 100$, $X_{j,t} = X_{j,t-1} + a_{j,t}$ $X_{j,t}' = X_{j,t} + kb_{j,t}$; $a_{j,t}, a_t, b_t, b_{j,t}$ sets of independent $N(0, 1)$ white noises. $k = 0$ gives random walks, $k = 1$ gives A.R.I.M.A. (0, 1, 1) series. H_0 = no relationship, is true. Series length = 50, number of simulations = 100, $\bar{R}^2$ = corrected R^2.

		Per cent times H_0 rejected[a]	Average Durbin–Watson d	Average $\bar{R}^2$	Per cent $\bar{R}^2 > 0.7$
			Random walks		
Levels	$m = 1$	76	0.32	0.26	5
	$m = 2$	78	0.46	0.34	8
	$m = 3$	93	0.55	0.46	25
	$m = 4$	95	0.74	0.55	34
	$m = 5$	96	0.88	0.59	37
Changes	$m = 1$	8	2.00	0.004	0
	$m = 2$	4	1.99	0.001	0
	$m = 3$	2	1.91	−0.007	0
	$m = 4$	10	2.01	0.006	0
	$m = 5$	6	1.99	0.012	0
			A.R.I.M.A. (0, 1, 1)		
Levels	$m = 1$	64	0.73	0.20	3
	$m = 2$	81	0.96	0.30	7
	$m = 3$	82	1.09	0.37	11
	$m = 4$	90	1.14	0.44	9
	$m = 5$	90	1.26	0.45	19
Changes	$m = 1$	8	2.58	0.003	0
	$m = 2$	12	2.57	0.01	0
	$m = 3$	7	2.53	0.005	0
	$m = 4$	9	2.53	0.025	0
	$m = 5$	13	2.54	0.027	0

[a]Test at 5% level, using an overall test on $\bar{R}^2$.

Source: Granger and Newbold (1974).

Spurious Regressions

- Two common responses:
  - Use HAC standard errors (e.g., Newey-West, 1987)
  - Series are persistent but don't have a unit root.

Spurious Regressions

- Two common responses:
  - Use HAC standard errors (e.g., Newey-West, 1987)
  - Series are persistent but don't have a unit root.
- Granger, Hyung, and Jeon (2001)

$$X_t = \alpha + \beta Y_t + u_t$$

$$X_t = \theta_x X_{t-1} + \epsilon_{x,t}$$

$$Y_t = \theta_y Y_{t-1} + \epsilon_{y,t}$$

Table 1. *Regressing between two independent AR series* ($\theta = \theta_x = \theta_y$), *percentage of* $|t| > 1.96$

Method	NOBS	$\theta = 0$	$\theta = 0.25$	$\theta = 0.5$	$\theta = 0.75$	$\theta = 0.9$	$\theta = 1.0$
OLS	100	5.3	7.6	13.3	29.1	51.5	77.0
	500	5.8	7.5	16.3	31.5	51.6	90.0
	2 000	5.8	7.1	13.5	29.4	52.5	94.5
	10 000	4.3	6.6	12.2	30.6	52.3	97.6
	∞	5.0	7.0	13.0	30.0	53.0	100.0
BART	100	7.6	7.7	9.9	16.5	30.6	62.0
	500	6.4	6.8	9.0	14.1	23.9	79.6
	2 000	6.0	5.9	6.1	10.3	16.3	86.4
	10 000	4.6	5.2	5.5	7.7	12.8	92.5
	∞	5.0	5.0	5.0	5.0	5.0	100.0

Notes: 1. The number of iteration = 1000.
2. % of rejection, i.e., absolute value of t-value > 1.96.
3. ∞ means asymptotic case.
4. To avoid the problem of fixing X_0 and Y_0, 100 pre-samples are generated and let $X_{-100} = Y_{-100} = 0$.
5. The number of rejections (BART) depends on the number of lags (l) used to calculate $\hat{v}$. $l =$ integer $[4(T/100)^{1/4}]$ is set.

Source: Granger, Hyung, and Jeon (2001).

Spurious Regressions

- Big problem even if series are stationary if they are pretty persistent and sample is small
- Newey-West standard errors have very bad small sample properties

Spurious Regressions

- Big problem even if series are stationary if they are pretty persistent and sample is small
- Newey-West standard errors have very bad small sample properties

- Accurate standard errors require more sophisticated methods
- Lazarus-Lewis-Stock 18 suggest improvements
- Even this not so good. No really satisfactory methods exist

Spurious Regressions

Why does this problem occur, intuitively?

Spurious Regressions

Why does this problem occur, intuitively?

- Observations are not independent!!
- Observations that are close in time are very correlated

Spurious Regressions

Why does this problem occur, intuitively?

- Observations are not independent!!
- Observations that are close in time are very correlated
- Intuitively, the key question is:
  - How many independent observations do I have?

    (With unit root, all observations are correlated)

Why does this problem occur, intuitively?

- Observations are not independent!!
- Observations that are close in time are very correlated
- Intuitively, the key question is:
  - How many independent observations do I have?
    (With unit root, all observations are correlated)
- Is higher frequency data useful?

Spurious Regressions

Why does this problem occur, intuitively?

- Observations are not independent!!
- Observations that are close in time are very correlated
- Intuitively, the key question is:
  - How many independent observations do I have?

    (With unit root, all observations are correlated)
- Is higher frequency data useful?
  - It does increase the number of data points
  - But the correlation between data points goes up
  - Intuitively: No new information about low frequency stuff

- Whether a sample is “small” or “large” is not so simple a question
- Depends on how correlated observations are
- You can have hundreds of thousands of observations but a “small sample” problem if correlation between observations is very high
- Cross-sections correlation can also be a problem (hence importance of “clustering” in constructing standard errors)

Rising Empirical Standards

1. "Revolution of identification"
   - More serious attention to credible identification of causal effects

2. Accurate standard errors
   - Clustering
   - Accounting for persistence

Has Meese-Rogoff 83 Stood the Test of Time?

Has Meese-Rogoff 83 Stood the Test of Time?

- Mostly yes!
- Rossi 13 provides comprehensive survey
- Mark 95 long-run predictability results most serious challenge
- See also more recent work on Taylor rule fundamentals (Molodtsova-Papell JIE 09)

Mark (1995)

- Simple monetary model:

$$e_t = f_t + c$$

$$f_t = (m_t - m_t^*) - \lambda(y_t - y_t^*)$$

- Even if monetary model doesn't work in the short run, it may work in the long run
- Estimates partial adjustment model:

$$e_{t+k} - e_t = \alpha_k + \beta_k(f_t - e_t) + \nu_{t+k,t}$$

Mark (1995)

$$e_{t+k} - e_t = \alpha_k + \beta_k(f_t - e_t) + \nu_{t+k,t}$$

- Sample period: 1973:2 - 1991:4
- Pseudo-out-of-sample period: 1981:4 - 1991:4
- Currencies: Canada, Germany, Switzerland, Japan
- Horizons: k = 1, 4, 8, 12, 16 (quarters)

Mark (1995): Econometric Issues

$$e_{t+k} - e_t = \alpha_k + \beta_k(f_t - e_t) + \nu_{t+k,t}$$

- Multiperiod forecasts induce correlation in error terms

Mark (1995): Econometric Issues

$$e_{t+k} - e_t = \alpha_k + \beta_k(f_t - e_t) + \nu_{t+k,t}$$

- Multiperiod forecasts induce correlation in error terms
- Stambaugh 86/99 bias
  - $f_t - e_t$ predetermined but not strictly exogenous
  - Past values of $e_{t+k} - e_t$ correlated with $f_t - e_t$
  - Causes finite sample bias in β_k

Mark (1995): Econometric Issues

$$e_{t+k} - e_t = \alpha_k + \beta_k(f_t - e_t) + \nu_{t+k,t}$$

- Multiperiod forecasts induce correlation in error terms
- Stambaugh 86/99 bias
  - $f_t - e_t$ predetermined but not strictly exogenous
  - Past values of $e_{t+k} - e_t$ correlated with $f_t - e_t$
  - Causes finite sample bias in β_k
- Standard errors produced using bootstrap that assumes $f_t - e_t$ follows AR(p)
  - But e_t and f_t may not be cointegrated
  - Small sample bias in estimating AR(p)

MARK (1995): ROBUSTNESS

- Why not use UK pound?

Mark (1995): Robustness

- Why not use UK pound?
- Mark calibrates $\lambda = 1$:

$$f_t = (m_t - m_t^*) - (y_t - y_t^*)$$

also no interest rate term. Why not estimate?

- Why not use UK pound?
- Mark calibrates $\lambda = 1$:

$$f_t = (m_t - m_t^*) - (y_t - y_t^*)$$

  also no interest rate term. Why not estimate?
- GNP for US, GDP for all other countries. Why?
- M3 for Canada, M1 for all other countries. Why?

TABLE 2—REGRESSION ESTIMATES AND BOOTSTRAP DISTRIBUTIONS

(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	(xi)	(xii)	(xiii)	(xiv)
k	$\hat{\beta}_k$	Adj-p	Adj-n	R^2	Adj-p	Adj-n	$t_k(20)$	MSL-p	MSL-n	$t_k(A)$	A	MSL-p	MSL-n
Canadian dollar:													
1	0.040	0.027	0.028	0.059	0.050	0.051	3.051	0.076	0.077	2.172	1	0.065	0.066
4	0.155	0.109	0.114	0.179	0.144	0.146	2.389	0.186	0.195	2.168	12	0.172	0.175
8	0.349	0.264	0.264	0.351	0.287	0.285	2.539	0.222	0.230	2.527	19	0.201	0.216
12	0.438	0.320	0.315	0.336	0.251	0.235	1.961	0.317	0.343	1.936	29	0.323	0.343
16	0.450	0.295	0.287	0.254	0.146	0.121	1.542	0.420	0.447	1.512	33	0.436	0.456
Deutsche mark:													
1	0.035	0.012	0.016	0.015	0.005	0.006	1.836	0.280	0.252	0.929	2	0.408	0.391
4	0.205	0.114	0.126	0.104	0.065	0.068	2.902	0.169	0.157	2.290	15	0.206	0.193
8	0.554	0.380	0.410	0.265	0.196	0.190	3.487	0.174	0.159	3.558	26	0.147	0.143
12	0.966	0.733	0.759	0.527	0.432	0.410	6.329	0.059	0.057	6.510	29	0.047	0.048
16	1.324	1.015	1.046	0.762	0.638	0.603	9.256	0.027	0.033	9.124	23	0.024	0.025
Swiss franc:													
1	0.074	0.046	0.046	0.051	0.042	0.042	2.681	0.109	0.119	2.073	2	0.084	0.086
4	0.285	0.171	0.171	0.180	0.147	0.145	3.248	0.121	0.126	3.196	14	0.096	0.102
8	0.568	0.356	0.350	0.336	0.278	0.276	4.770	0.080	0.085	4.696	21	0.078	0.073
12	0.837	0.527	0.519	0.538	0.458	0.452	8.013	0.026	0.024	8.013	20	0.026	0.021
16	1.086	0.706	0.671	0.771	0.673	0.655	17.406	0.001	0.002	12.665	14	0.005	0.006
Yen:													
1	0.047	0.014	0.016	0.020	0.010	0.011	1.396	0.388	0.365	1.331	3	0.285	0.259
4	0.263	0.136	0.138	0.125	0.088	0.090	2.254	0.271	0.262	2.153	14	0.247	0.231
8	0.575	0.328	0.329	0.301	0.233	0.232	3.516	0.199	0.189	3.496	19	0.188	0.177
12	0.945	0.592	0.579	0.532	0.432	0.427	4.889	0.129	0.143	4.735	17	0.153	0.156
16	1.273	0.819	0.802	0.694	0.565	0.548	4.919	0.154	0.156	4.901	16	0.174	0.177

Notes: The table presents OLS estimates of the regression $e_{t+k} - e_t = \alpha_k + \beta_k(f_t - e_t) + \nu_{t+k,k}$, where $f_t \equiv (m_t - m_t^*) - (y_t - y_t^*)$. The (Gaussian) parametric and nonparametric bootstrap distributions are generated under the null hypothesis that the regressor follows an AR(4) for the Canadian dollar, the Swiss franc, and the yen, and an AR(5) for the deutsche mark. Exchange rates are dollars per unit of foreign currency. Adj-p and Adj-n are bias-adjusted values obtained by subtracting median values generated by the parametric and nonparametric bootstrap distributions, respectively, from the estimates. MSL-p and MSL-n are, respectively, the parametric and nonparametric bootstrap marginal significance levels for a one-tail test. A is the truncation lag determined by Andrews's (1991) univariate AR(1) rule used for constructing the t ratios with the data.

Source: Mark (1995). Note: Big β, big R^2, large $t_k(20)$ for DM, CHF, JPY

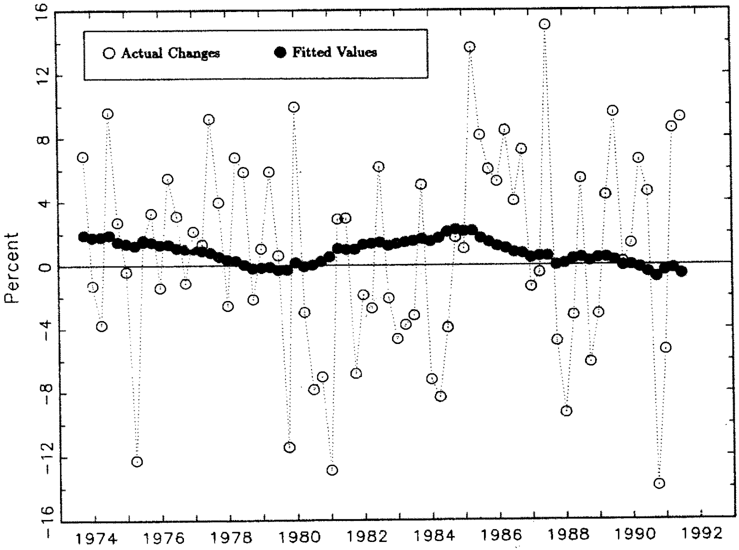


FIGURE 1. ONE-QUARTER CHANGES IN THE LOG DOLLAR/DEUTSCHE-MARK EXCHANGE RATE

Source: Mark (1995)

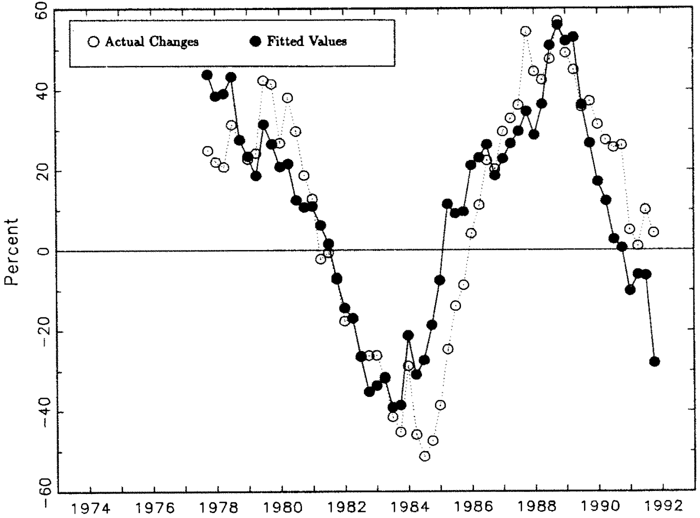


FIGURE 5. SIXTEEN-QUARTER CHANGES IN THE LOG DOLLAR/DEUTSCHE-MARK EXCHANGE RATE

Source: Mark (1995)

TABLE 4—OUT-OF-SAMPLE FORECAST EVALUATION

(i) k	(ii) IN/RW	(iii) OUT/IN	(iv) OUT/RW	(v) MSL-p	(vi) MSL-n	(vii) $\mathcal{DM}(20)$	(viii) MSL-p	(ix) MSL-n	(x) $\mathcal{DM}$(A)	(xi) A	(xii) MSL-p	(xiii) MSL-n
Canadian dollar:												
1	0.960	1.040	0.998	0.209	0.194	0.061	0.215	0.202	0.036	1	0.218	0.201
4	0.889	1.258	1.119	0.571	0.538	−1.270	0.526	0.487	−0.925	8	0.494	0.468
8	0.675	1.695	1.145	0.447	0.397	−1.036	0.427	0.377	−0.890	17	0.420	0.390
12	0.654	2.197	1.436	0.613	0.578	−1.916	0.574	0.556	−1.661	18	0.587	0.579
16	0.799	2.128	1.699	0.654	0.636	−2.596	0.578	0.542	−1.857	15	0.567	0.555
Deutsche mark:												
1	0.988	1.027	1.015	0.397	0.339	−0.932	0.458	0.393	−0.846	4	0.536	0.493
4	0.927	1.120	1.037	0.345	0.288	−1.345	0.563	0.511	−0.852	9	0.478	0.427
8	0.833	1.203	1.002	0.268	0.217	−0.027	0.270	0.220	−0.020	18	0.270	0.221
12	0.670	1.188	0.796	0.127	0.092	4.246	0.068	0.059	0.094	16	0.151	0.136
16	0.431	1.216	0.524	0.040	0.025	8.719[a]	0.061	0.047	8.719	18	0.021	0.011
Swiss franc:												
1	0.972	1.026	0.997	0.305	0.266	0.066	0.320	0.278	0.064	3	0.315	0.271
4	0.886	1.108	0.981	0.291	0.263	0.218	0.304	0.272	0.162	12	0.298	0.274
8	0.780	1.176	0.917	0.256	0.219	0.703	0.260	0.236	0.560	17	0.253	0.227
12	0.625	1.181	0.738	0.152	0.132	2.933	0.161	0.137	0.938	13	0.255	0.211
16	0.335	1.229	0.411	0.033	0.023	9.650[b]	0.080	0.058	1.996	8	0.192	0.159
Yen:												
1	0.962	1.027	0.988	0.304	0.257	1.571	0.168	0.132	0.836	3	0.177	0.134
4	0.822	1.129	0.928	0.257	0.207	2.302	0.151	0.118	1.487	10	0.134	0.105
8	0.688	1.191	0.819	0.214	0.162	3.096	0.142	0.117	1.803	13	0.152	0.117
12	0.536	1.329	0.712	0.196	0.148	3.319	0.174	0.148	1.147	17	0.164	0.135
16	0.363	1.579	0.574	0.152	0.119	5.126	0.178	0.160	3.096	16	0.151	0.131

Notes: The table presents ratios of root-mean-squared errors for the regression's out-of-sample forecasts (OUT), the driftless random walk (RW), and the in-sample regression residual during the forecast period (IN). The first forecast is made on 1981:4. $\mathcal{DM}(20)$ and $\mathcal{DM}$(A) are the Diebold-Mariano statistics constructed using the method of Newey and West (1987) with the truncation lag of the Bartlett window set to 20 and set by Andrews's (1991) AR(1) rule, respectively. In instances where the estimated spectral density at frequency zero of the squared error differential is nonpositive (see footnote 8), the Bartlett-window truncation lag is decreased by 1. MSL-p and MSL-n are marginal significance levels, generated by the parametric and nonparametric bootstrap distributions, respectively, for one-tail tests.

Source: Mark (1995). Note: OUT/RW much smaller than 1.

True Out-of-Sample Test

- Jon wrote a class paper on this for Jim Stock's Time Series class in 2003
- True out-of-sample period: 1992:1-2000:4

True Out-of-Sample Test

- Jon wrote a class paper on this for Jim Stock's Time Series class in 2003
- True out-of-sample period: 1992:1-2000:4
- Used slightly different data:
  - M2 as opposed to M3 for Canada
  - GDP as opposed to GNP for US
  - Results sensitive to this (not comforting)
- Main results do not survive in 1990s

Table 1 -- Replication, Extentions and Out of Sample Performance

	Mark's published results		Mark's data -- My replication		Current Data -- Mark's sample period		Mark's RMSE results	RMSE ratios for 1990's
	beta	R2	beta	R2	beta	R2	OUT/RW	OUT/RW
	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Canadian dollar vs. U.S. dollar								
1	0.040	0.059	0.041	0.061	-0.005	0.002	0.998	0.963
4	0.155	0.179	0.157	0.182	-0.041	0.022	1.119	0.933
8	0.349	0.351	0.350	0.354	-0.073	0.020	1.145	0.932
12	0.438	0.336	0.442	0.342	-0.187	0.061	1.436	1.037
16	0.450	0.254	0.456	0.262	-0.305	0.107	1.699	1.208
Deutsche mark vs. U.S. dollar								
1	0.035	0.015	0.037	0.016	0.036	0.016	1.015	1.029
4	0.205	0.104	0.204	0.104	0.181	0.087	1.037	0.987
8	0.554	0.265	0.552	0.264	0.503	0.231	1.002	0.992
12	0.966	0.527	0.961	0.524	0.911	0.485	0.796	1.511
16	1.324	0.762	1.318	0.758	1.274	0.715	0.524	1.957
Swiss franc vs. U.S. dollar								
1	0.074	0.051	0.073	0.050	0.087	0.073	0.997	0.984
4	0.285	0.180	0.284	0.178	0.314	0.227	0.981	0.937
8	0.568	0.336	0.566	0.335	0.571	0.351	0.917	0.786
12	0.837	0.538	0.834	0.536	0.804	0.511	0.738	0.763
16	1.086	0.771	1.085	0.770	1.064	0.751	0.411	0.980
Japanese Yen vs. U.S. dollar								
1	0.047	0.020	0.047	0.020	0.030	0.011	0.988	1.014
4	0.263	0.125	0.264	0.126	0.195	0.085	0.928	1.117
8	0.575	0.301	0.576	0.302	0.452	0.227	0.819	1.177
12	0.945	0.532	0.948	0.534	0.769	0.421	0.712	1.139
16	1.273	0.694	1.274	0.696	1.063	0.557	0.574	1.185

Source: Steinsson (2003)

Follow Up on Mark (1995)

- Killian 99 makes same point as I did. Also critiques bootstrap.
- Faust-Rogers-Wright 03: Doesn't work with other vintages of data

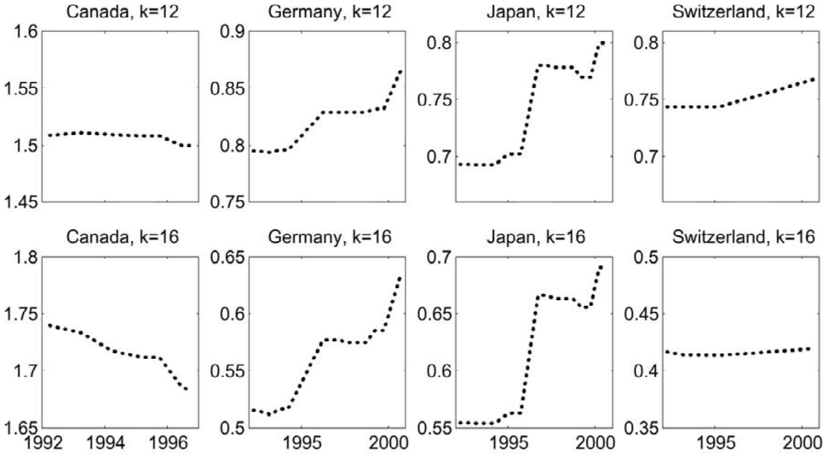


Fig. 6. Out-of-sample relative RMSE using different data vintages (Mark's sample period).

Source: Faust-Rogers-Wright (2003)

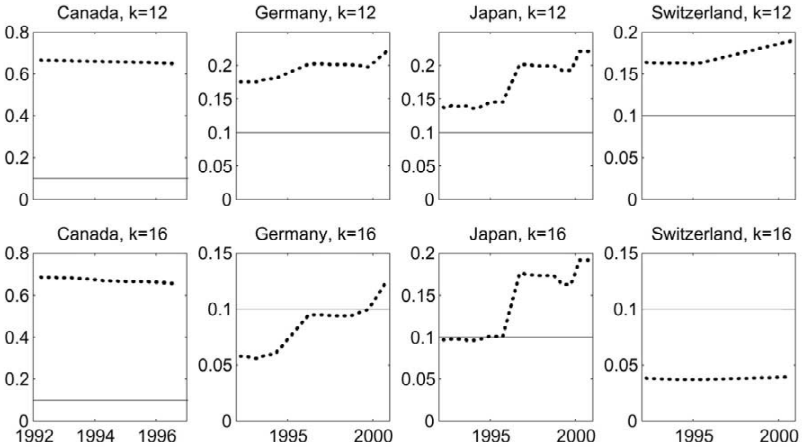


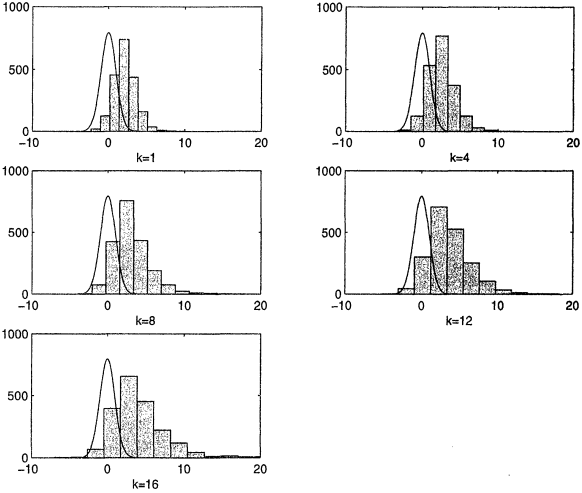
Fig. 7. Bootstrap *P*-values for out-of-sample relative RMSE using different data vintages (Mark's sample period).

Source: Faust-Rogers-Wright (2003)

Follow Up on Mark (1995)

- Killian 99 makes same point as I did. Also critiques bootstrap.
- Faust-Rogers-Wright 03: Doesn't work with other vintages of data
- Berkowitz-Giorgianni 01:
  - Mark's bootstrap assumes e_t and f_t are cointegrated (It assume AR(p) for $z_t = f_t - e_t$)
  - Standard errors much larger under null of no-cointegration (also standard diagnostic test stats fail in this case)

FIGURE 1.—MONTE CARLO DISTRIBUTIONS OF NEWEY-WEST T-STATISTICS UNDER H_0: INDEPENDENCE



For each horizon of interest, $k = 1, 4, 8, 12$, and 16, we plot the histogram of the Newey-West t-statistic with a bandwidth of 20. The solid line corresponds to the density of a t-distributed random variable with degrees of freedom equal to $85 - (k + 1)$.

Source: Berkowitz-Giorgianni (2001)

TABLE 1.—LONG-HORIZON MONTE CARLO ESTIMATES NULL HYPOTHESIS: INDEPENDENCE
(SPOT RATES FOLLOW A RANDOM WALK AND FUNDAMENTALS AN AR(2) PROCESS)

k	%-ile	$\hat{\beta}_k$	t(LS)	t(20)	t(A)	R_k^2	OUT/RW	DM(20)	DM(A)
Panel A: sample size = 85									
1	50	0.046	1.479	2.028	1.525	0.026	1.008	−0.377	−0.300
	90	0.112	2.517	3.941	2.740	0.072	0.969	2.222	1.341
	95	0.143	2.827	4.502	3.132	0.088	0.957	3.197	1.758
4	50	0.183	2.982	2.359	2.092	0.101	1.032	−0.441	−0.372
	90	0.403	5.220	4.823	4.180	0.256	0.881	2.458	1.816
	95	0.495	6.012	5.666	4.947	0.314	0.840	3.430	2.400
8	50	0.354	4.270	2.769	2.635	0.195	1.050	−0.420	−0.366
	90	0.711	7.741	6.079	5.689	0.444	0.768	2.752	2.230
	95	0.839	9.004	7.442	6.813	0.519	0.699	4.058	3.015
12	50	0.507	5.333	3.200	3.203	0.286	1.045	−0.320	−0.301
	90	0.958	9.798	7.173	6.839	0.574	0.678	3.108	2.574
	95	1.067	11.267	8.928	8.112	0.641	0.605	4.552	3.524
16	50	0.635	6.132	3.465	3.585	0.359	1.029	−0.223	−0.194
	90	1.122	11.505	8.292	7.976	0.663	0.601	3.796	2.956
	95	1.244	13.360	9.790	9.488	0.727	0.516	5.089	4.122
Panel B: sample size = 1085									
1	50	0.004	1.559	1.593	1.572	0.002	0.999	0.198	0.134
	90	0.011	2.565	2.734	2.595	0.006	0.991	2.495	1.559
	95	0.013	2.834	2.991	2.854	0.007	0.987	3.107	1.943
4	50	0.016	3.102	1.651	1.646	0.009	0.997	0.212	0.150
	90	0.041	5.145	2.805	2.853	0.024	0.962	2.562	2.014
	95	0.050	5.683	3.142	3.160	0.029	0.946	3.423	2.571
8	50	0.031	4.433	1.728	1.723	0.017	0.995	0.181	0.164
	90	0.082	7.385	2.976	3.053	0.048	0.920	2.908	2.452
	95	0.100	8.088	3.321	3.391	0.057	0.890	3.915	3.157
12	50	0.046	5.449	1.815	1.791	0.026	0.992	0.208	0.189
	90	0.121	9.098	3.146	3.191	0.072	0.874	3.509	3.007
	95	0.149	9.982	3.520	3.614	0.085	0.831	4.884	3.951
16	50	0.062	6.260	1.911	1.835	0.035	0.992	0.173	0.149
	90	0.158	10.505	3.292	3.374	0.093	0.826	4.433	3.807
	95	0.198	11.504	3.715	3.815	0.110	0.773	6.308	5.278

The table presents estimated slope coefficients, $\hat{\beta}_k$, for equation (2) with the LS t-statistics, heteroskedasticity, and autocorrelation-corrected t-statistics using a Bartlett kernel and a truncation lag of 20 and Andrews' (1991) rule: respectively, t(LS), t(20), and t(A). OUT/RW denotes the ratio of regression mean-squared out-of-sample forecast error to the random-walk, mean-squared, out-of-sample forecast error. DM(20) and DM(A) denote the Diebold-Mariano statistics with a Bartlett kernel and truncation lags of 20 and via Andrews' (1991) rule, respectively.

Source: Berkowitz-Giorgianni (2001)

Table 4.—Long-Horizon Regression Estimates Null Hypothesis: No Cointegration (1973:2–1991:4)

	k	$\hat{\beta}_k$	$t(20)$	p-val	$t(A)$	p-val	R_k^2	OUT/RW	p-val	DM(20)	p-val	DM(A)	p-val
Canadian Dollar	1	0.040	3.051	0.095	2.172	0.086	0.059	0.998	0.405	0.061	0.441	0.036	0.443
	4	0.155	2.398	0.217	2.168	0.190	0.179	1.119	0.412	−1.270	0.849	−0.925	0.975
	8	0.349	2.539	0.225	2.527	0.212	0.351	1.145	0.712	−1.036	0.958	−0.890	0.901
	12	0.438	1.961	0.352	1.936	0.350	0.336	1.436	0.317	−1.916	0.592	−1.661	0.695
	16	0.450	1.542	0.458	1.512	0.465	0.254	1.699	0.196	−2.596	0.466	−1.857	0.620
German Mark	1	0.035	1.836	0.510	0.929	0.668	0.015	1.015	0.969	−0.932	0.724	−0.846	0.669
	4	0.205	2.902	0.354	2.290	0.406	0.104	1.037	0.914	−1.345	0.522	−0.852	0.780
	8	0.554	3.487	0.354	3.558	0.313	0.265	1.002	0.809	−0.027	0.814	−0.020	0.814
	12	0.966	6.329	*0.165	6.510	*0.135	0.527	0.796	*0.406	4.246	0.093	0.094	0.563
	16	1.324	9.256	0.096	9.124	0.082	0.762	0.524	*0.113	8.719	0.030	8.719	0.016
Japanese Yen	1	0.047	1.396	0.516	1.331	0.423	0.020	0.988	0.477	1.571	0.286	0.836	0.322
	4	0.263	2.254	0.353	2.153	0.341	0.125	0.928	0.396	2.302	0.215	1.487	0.310
	8	0.575	3.516	0.228	3.496	0.227	0.301	0.819	0.304	3.096	0.172	1.803	0.309
	12	0.945	4.889	0.166	4.735	0.190	0.532	0.712	0.233	3.319	0.173	1.147	0.276
	16	1.273	4.919	0.216	4.901	0.199	0.694	0.574	0.142	5.126	0.109	3.096	0.241
Swiss Franc	1	0.074	2.681	0.210	2.073	*0.166	0.051	0.997	0.642	0.066	0.704	0.064	0.693
	4	0.285	3.248	0.181	3.196	*0.131	0.180	0.981	0.596	0.218	0.676	0.162	0.686
	8	0.568	4.770	0.095	4.696	0.081	0.336	0.917	0.458	0.703	0.621	0.560	0.635
	12	0.837	8.013	0.032	8.013	0.023	0.538	0.738	0.214	2.933	0.203	0.938	0.624
	16	1.086	17.41	0.006	12.66	0.010	0.771	0.411	0.026	9.650	#0.019	1.996	0.411

	ρ_1	ρ_2	ρ_3	ρ_4	σ^2
Canadian Dollar	1.227	0.028	−0.045	−0.233	0.011
German Mark	1.253	−0.305			0.016
Japanese Yen	1.267	−0.062	−0.204		0.013
Swiss Franc	1.916	−1.154	0.236		0.014

Data were kindly supplied by Nelson Mark (originally taken from the OECD *Main Economic Indicators*). For detailed descriptions of the statistics, see the notes to table 1.
* indicates p-values that are no longer significant at a 90% level, but were under the null of cointegration.
\# indicates the reverse.
The table presents least-squares estimates of equation (2) over horizons of k = 1, 4, 8, 12, and 16 quarters and Monte Carlo p-values, tabulated under the null hypothesis that s_t is a random walk, independent of f_t. The f_t are generated using the following AR models with lag order selected by the BIC criterion.

Source: Berkowitz-Giorgianni (2001)