Excerpts – Predictors of Youth Literacy in Thailand. A Time Series Regression Analysis. Excerpts. Full 50 page paper available upon request.

Predictors of Youth Literacy in Thailand: A Time Series Model

Stephen Chemsak

Teachers College, Columbia University

Analysis of Political Data – W4911

Professor Shapiro

Columbia University School of International and Public Affairs

Dataset

Year	Urban population (% of total)	GNI per capita, Atlas method (current $US)	Aid (% of GNI)	Age dependency ratio (dependents to working-age population)	Literacy rate, youth total (% of people ages 15-24)
1970	21	210	1.0458	0.9179589	94.13091
1971	21	210	0.8522	0.90268873	94.42764
1972	22	220	0.6608	0.88778154	94.72272
1973	23	270	0.5702	0.87307914	95.01715
1974	23	330	0.5255	0.85862372	95.31302
1975	24	390	0.5852	0.84449003	95.61095
1976	24	430	0.9986	0.82217825	95.80374
1977	25	460	0.6725	0.80056425	95.99789
1978	26	540	1.0919	0.779543	96.19346
1979	26	620	1.4508	0.75907986	96.3901
1980	27	730	1.3039	0.73921299	96.58737
1981	27	780	1.1858	0.71746433	96.75976
1982	27	780	1.0794	0.69651619	96.93267
1983	28	780	1.0759	0.67618853	97.10534
1984	28	810	1.1184	0.65649334	97.27654
1985	28	820	1.1988	0.63740556	97.44514
1986	28	860	1.1017	0.61993206	97.58192
1987	29	980	0.9461	0.60291798	97.71788
1988	29	1200	0.8793	0.58652455	97.85451
1989	29	1370	1.0134	0.57062796	97.99309
1990	29	1540	0.9486	0.55525094	98.13403
1991	30	1720	0.7408	0.5424036	98.2199
1992	30	1940	0.6794	0.52995417	98.30672
1993	30	2170	0.4711	0.51790583	98.39394
1994	30	2430	0.4065	0.50632714	98.48095
1995	30	2780	0.5083	0.49511492	98.56729
1996	30	3000	0.4651	0.48062483	98.63455
1997	31	2760	0.426	0.46709355	98.70134
1998	31	2090	0.6522	0.45462068	98.76832
1999	31	1980	0.8524	0.44303343	98.83598
2000	31	1990	0.5781	0.43243307	97.97539

Scatterplots for Impure Serial Correlation/Functional Form

```
. graph twoway scatter lit gnipercap
```
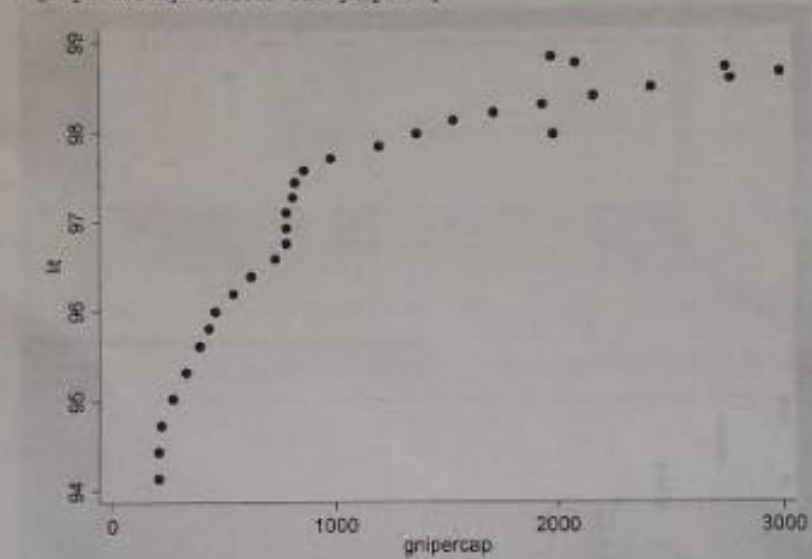

Suggests impure serial correlation/incorrect functional form so log was taken

```
. graph twoway scatter lit gnipercaplog
```

Time Series Regression using Prais-Winsten, Cochrane-Orcutt, and Newey-West Standard Errors

Because the Durbin-Watson test for serial correlation was inconclusive in the original model, Prais- Winsten,Cochrane-Orcutt, and Newey-West Standard Errors regressions were run to correct for any possible serial correlation.

```
. prais til gnipercap urbanpop aidofgni adr


Prais-Winsten AR(1) regression -- iterated estimates
```

Source	SS	df	MS
Model	214.738894	4	53.6847236
Residual	.131873435	26	.005072055
Total	214.870768	30	7.16235892

Number of obs = 31
$F_{(4, 26)}$ = 10584.41
Prob > F = 0.0000
R-squared = 0.9994
Adj R-squared = 0.9993
Root MSE = .07122

til	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
gnipercap	.0000152	.000076	0.20	0.843	-.000141	.0001715
urbanpop	.0413144	.0292335	1.41	0.169	-.018776	.1014048
aidofgni	-.0238866	.0741804	-0.32	0.750	-.1763666	.1285934
adr	-8.849176	.9820025	-9.01	0.000	-10.86771	-6.830641
cons	101.4974	1.384228	73.32	0.000	98.6521	104.3427

rho	.9843789

```
Durbin-Watson statistic (original)      .328138
Durbin-Watson statistic (transformed)  0.610590
```

The Durbin-Watson test statistic is .611. For n=31 and k=4, the d_L=1.16 and the d_U=1.74. Therefore, there is evidence of serial correlation.