TA session# 8

Jun Sakamoto

November 29,2018

Contents

We will look at some empirical studies for panel data analysis.

1 Empirical study 1

Z. Lv and R. Yang (2018) "Does women's participation in politics increase female labor participation? Evidence from panel data analysis" *Economic Letters* 170, 35-38.

Their reserch question as follows. Whether women's participation in politics affects female labor participation rates? They used the following equation for this verification.

$$FLRP^{*}_{it} = \alpha + \beta WPP_{it} + \gamma CV_{it} + \mu_{it} + \epsilon_{it}$$
$$where FLRP^{*}_{it} = \log\left(\frac{FLRP_{it}}{1 - FLRP_{it}}\right) \quad (1)$$

$FLRP_{it}$: Female labor participation rates.
WPP_{it}: Women ' s political participation index by Sundstrom (2017). CV_{it}: Set of control variables.

They use a sample of 99 countries with data covering the years from 1991 and 2012. $FLPR$ is defined as the number of female labor participants of age 1564 divided by the total female population of the same age group (1564), and labor force participation is defined as employed (paid and unpaid family workers) plus unemployed (actively seeking work).

Column 1shows the results from simple static fixed-effects specifications. WPP has a positive and statistically significant effect on FLPR at the 1% level. Also 4th column show the result of dynamic panel data by GMM system estimator. if FLPR is persistent or/and there exist the reverse causality between FLPR and WPP, then the results obtained from the static model could be biased. To take into account the dynamic effects and the endogeneity issue, we further apply the dynamic panel data to study the impact of WPP on FLPR. Specifically, they adopt the system GMM estimator.

Table 2
Main results.

Variable	(1)	(2)	(3)	(4)
FLPR*(lagged)				0.9243***(32.74)
WPP	0.1682***(4.59)		0.0357*(1.77)	0.0747**(2.30)
WNP		0.3594***(4.50)		
GDP per capita	−1.8387***(−13.48)	−2.3232***(−10.36)		−0.2995*(−1.69)
GDP per capita Sq.	0.1112***(14.25)	0.1379***(10.95)		0.0171*(1.79)
Fertility	−0.0261**(−2.42)	−0.0521***(−3.06)	−0.0113(−0.59)	0.0095(0.77)
Urbanization	0.0054***(3.86)	0.0147***(7.89)	0.0124**(2.83)	0.0000(0.15)
Unemployment rates	−0.0017(−1.15)	0.0032**(1.99)	−0.0003(−0.37)	−0.0017*(−1.83)
Education	0.0003(0.81)	−0.0002(−0.39)	−0.0000(−0.11)	0.0005(1.58)
Constant	7.3957***(12.25)	9.0237***(8.79)		1.2060(1.42)
Number of cross-sections	99	60	99	99
Number of instruments				90
AR(1) test *p*-value				0.000
AR(2) test *p*-value				0.483
Hansen J test *p*-value				0.460

Notes: Here WPP means the V-Dem women's political participation index developed by Sundström et al. (2017), while WNP represents the proportion of seats held by women in national parliaments. Figures in parentheses are *t*-values. Instruments are collapsed as suggested by Roodman (2009). Hansen test is a test for overidentifying restrictions, the null hypothesis is that the instruments are valid. AB test (1), (2) are Arellano–Bond test for AR(1), AR(2) in first differences, respectively, the null hypothesis for (1) is that the first-differenced regression errors show no first-order serial correlation, the null hypothesis for (2) is that the first-differenced regression errors show no second-order serial correlation.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

2 Empirical study 2

E. S. Mayfield and R. G. Murphy (1992) "Interest rate parity and the exchange risk premium Evidence from panel data' ' *Economic Letters* 40, 319-324.

This paper provides evidence that a time-varying risk premium is responsible for the rejection of the interest rate parity theory. The interest rate parity theory in linearized form, modified to include a time-varying risk premium, states that the expected change in the domestic price of foreign exchange can be related to the difference between domestic and foreign interest rates:

$$E_t S^i_{t+n} - S^i_t = \rho^i_{n,t} + [R_{n,t} - R^i_{n,t}] \tag{2}$$

S^i_t is the logarithm of the domestic price of foreign currency i in period $t+n$, $R_{n,t}$ and $R^i_{n,t}$, are n-period interest rates on similar assets denominated in domestic currency and foreign currency i and $\rho^i_{n,t}$ is the time varying risk premium. Under the assumption of rational expectations, the actual exchange rate will be equal to the expected exchange rate plus a white noise forecast error:

$$S^i_{t+n} = E_t S^i_{t+n} + \gamma^i_{t+n} \tag{3}$$

From these equations, we can get

$$S^i_{t+n} - S^i_t = \rho^i_{n,t} + [R_{n,t} - R^i_{n,t}] + \gamma^i_{t+n}. \tag{4}$$

They use following equation to estimate.

$$S^i_{t+n} - S^i_t = \phi + \eta_t + \mu_i + \alpha[R_{n,t} - R^i_{n,t}] + \gamma^i_{t+n} \tag{5}$$

Table 1
Estimates with common coefficient. [a,b,c,d]

	Three-month			Six-month		
	(1)	(2)	(3)	(4)	(5)	(6)
α	−0.2460 (0.4608)	−1.4036 (0.6763)	0.7929 (0.2864)	−0.1667 (0.4871)	−1.4923 (0.7255)	0.5421 (0.2313)
$\overline{R}^2$	−0.0001	0.0261	0.0299	−0.0005	0.0433	0.0230
DF	559	557	370	550	548	364
t-stat	−2.704	−3.554	−0.723	−2.395	−3.435	−1.980
p-value	0.0069	0.0004	0.4697	0.0166	0.0006	0.0477

[a] Columns (1) and (4) present results using OLS on the stacked data set. Columns (2) and (5) present results allowing for fixed currency effects. Columns (3) and (6) present results allowing for fixed currency and fixed time effects.
[b] Asymptotic standard errors, adjusted for the moving average process implied by overlapping observations intervals, are in parentheses.
[c] P-value is the two-sided probability of obtaining a value at least as large as the reported t-statistic under the null hypothesis that $\alpha = 1$.
[d] The data are end-of-month from January 1975 to October 1990. This gives 190 observations for each currency. The estimates for the 3-month horizon, however, require a lag of 3 months, so there are 187 useable data points. Likewise, the estimates for the 6-month horizon require a lag of 6 months, so there are 184 useable data points. When we stack the data for the three currencies, the total number of data points are 561 for the 3-month horizon and 552 for the 6-month horizon.

Data for maturities at three and six months over four currency denominations (U.S. dollar, French franc, Swiss franc, and German mark) are employed. s The exchange rate is defined as the dollar price of the respective foreign currency. The data are for the last trading day of the month over the period January 1975 to October 1990.

Column (1) and (4) are assumed that the fixed currency and fixed time effects are zero ($\eta_t = \mu_i = 0$), and only allow for the presence of a constant term. Columns (2) and (5) report estimates assuming only a fixed currency effect $\eta_t = 0$ and Columns (3) and (6) provide estimates allowing for both fixed currency and fixed time effects. There results suggest that accounting for these unobserved but related movements in risk premia improves the ability of the interest rate parity theory to explain fluctuations in exchange rates. References

3 Empirical study 3

C.W. Hansen (2012) “The relation between wealth and health: Evidence from a world panel of countries” *Economic Letters* 115, 175-176.

This paper presents panel data evidence that documents a U-shaped relation between GDP per capita (wealth) and life expectancy (health). The basic empirical specification is given by the following reduced form relationship between wealth and health.

$$LGDP_{it} = \alpha_1 LLE_{it} + \alpha_2 LLE_{it}^2 + \delta_i + \mu_t + \nu_{it} \qquad (6)$$

where $LGDP_{it}$ is the level of wealth for country i in period t ,measured by the log of real GDP per capita, LLE_{it} is the level of health which is measured by log life expectancy at birth, δ_i and μ_t are time and country specific effects, and ν is the disturbance term. In all the regressions reported in this paper δ_i and μ_t are removed from the disturbance term by fixed effect estimation and inclusion of time dummies. The panel data set constructed for the purpose of estimating above equation is based on 10-yearly observations from 1940 to 1980 with data from Acemoglu and Johnson (2007).

To account for dynamic effects and the endogeneity problem, I also estimate the following equation:

$$LGDP_{it} = \gamma LGDP_{it-1} + \beta_1 LLE_{it} + \beta_2 LLE_{it}^2 + \delta_i + \mu_t + \nu_{it}. \qquad (7)$$

They estimate β_1 and β_2 by GMM estimation as suggested by Arellano and Bond (1991).

Table 1
The relation between wealth and health, 1940–1980.

	Log GDP per capita, *LGDP*			
	(1)	(2)	(3)	(4)
	FE (within)	FE (within)	GMM (Arellano–Bond)	GMM (Arellano–Bond)
LLE	−0.42 (0.27)	−19.60*** (2.62)	−0.01 (0.16)	−9.46*** (3.53)
LLE^2		2.58*** (0.35)		1.26*** (0.47)
No. of countries	119	119	119	119
Turning point	–	45.16	–	43.23
Sargan (*p*-value)	–	–	0.12	0.22

Notes: Standard errors (SE) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. SE in (1) and (2) are cluster robust. All regressions include time fixed effects.

They find that α_1 is negative and insignificantly different from zero, which might mistakenly lead to the conclusion that wealth and health are not related after all. However, this conclusion is reversed when allowing for a nonmonotonic relationship. In particular, column (2) shows a negative α_1 and a positive α_2 -both statistically significant at the 1% levelwhich reveals that wealth follows a U-shaped path in the country level of health. This economical interpret is as follow. In an early stage of development, the effect of health improvements on wealth is negative because, at this stage, the only effect is to increase the size of the population which possibly has an adverse effect on wealth. In contrast, at a later stage of development, health improvements may induce human capital skills and the so-called Malthusian population link may be broken so that health improvements actually lead to a lower population size.

4 Empirical study 4

J. Wildman and B. Hollingsworth "Public smoking bans and self-assessed health: Evidence from Great Britain" *Economic Letters* 118, 209-212.

This paper investigates the impact of a public smoking ban on self-reported health status in Great Britain. They use the implementation of public smoking bans in England and Scotland as a natural experiment to estimate the causal effect of smoking bans on self-reported health. Scotland introduced a smoking ban in March

2006 (England followed in July 2007). This time difference is used to provide treatment (Scotland) and control (England) groups. The estimating equation is:

$$Health_{it} = \alpha_i + \beta_1 T_t + \beta_2 SB_i + \tau T \cdot SB_{it} + X_{it}\gamma + \nu_{it} \tag{8}$$

where $Health_{it}$ is the outcome of interest, SB_i demonstrates whether an individual was exposed to a smoking ban (in this case residing in Scotland), T_t is a time effect common to both groups, X_{it} is a matrix of control variables. α_i is an individual fixed effect and ν is the idiosyncratic error term. The treatment effect is identified as the parameter τ.

They use data from England and Scotland from the British Household Panel Survey (BHPS), a panel survey which has run since 1991. The introduction of the smoking ban in Scotland in 2006 coincides with wave 16 of the BHPS. Wave 1 consisted of over 5000 households providing around 10,000 individual interviews from England and Scotland.

Their outcome measure is self-assessed health (SAH). In all waves (except wave 9 (1999)) individuals were asked "Please think back over the last 12 months about how your health has been. Compared to people of your own age, would you say that your health has on the whole been ⋯Excellent, Good, Fair, Poor or Very Poor".2 This variable is dichotomised into a variable indicating good health (respondents answering excellent or good).

Table 1
Summary statistics: pooled means and standard deviations from 16 waves of data.[b]

		Mean	Std. dev.
SAH good	Dichotomised self-assessed health = 1 if answered good or excellent to "Health status over last 12 months" (incl wave 9 responses)	0.74	0.44
Limiting condition	Health condition limits activity (14 waves of data included, $N = 23,170$)	0.08	0.27
Treatment[a]	Resides in Scotland following smoking ban	0.09	0.07
False treatment[a]	Resides in Scotland in the period preceding the smoking ban	0.09	0.07
Women	Gender	0.46	0.50
Age	Age at 1st December in interview wave	40.23	8.71
Household size	Number of people in household	3.12	1.28
Number of children	Number of children less than 16 in household	0.83	1.04
Never married	Marital status = 1 if never married, 0 otherwise	0.13	0.34
Married/couple	Marital status = 1 if married/living together as couple, 0 otherwise	0.78	0.41
Widowed	Marital status = 1 if widowed, 0 otherwise	0.01	0.09
Divorced/separated	Marital status = 1 if divorced/separated, 0 otherwise	0.08	0.27
Log income	Log of annual household income	10.19	0.66
Employed	Job status = 1 if employed, 0 otherwise	0.77	0.42
Self employed	Job status = 1 if self employed, 0 otherwise	0.11	0.31
Unemployed	Job status = 1 if unemployed, 0 otherwise	0.03	0.17
Retired	Job status = 1 if retired, 0 otherwise	0.01	0.10
Family carer	Job status = 1 if family carer, 0 otherwise	0.04	0.20
Studying	Job status = 1 if studying, 0 otherwise	0.01	0.09
Disabled	Job status = 1 if disabled, 0 otherwise	0.03	0.18
Government training	Job status = 1 if on Government training scheme, 0 otherwise	0.002	0.04
Other job status	Job status = 1 if other job status, 0 otherwise	0.003	0.06
Scotland	Country of residence = 1 if Scotland, 0 if England	0.08	0.28

[a] Proportion of sample affected by treatment $N = 26,480$.
[b] The data comes from the BHPS waves 1–16. Questions were asked, in the same form, in every wave of the survey. Exceptions are the two health questions. As outlined in the text, the question forming the basis of the dependent variable (SAH good) was asked in a different form in wave 9, and the health condition limits activities question (limiting condition) was not asked in waves 9 and 14. The treatment variables were created using the region of residence and the date of interview.

Table 2
Treatment effects.

	Model 1 All	Model 2 Never smoked = 1	Model 3 Never smoked = 0
SAH			
Treatment effect	0.033	**0.096****	−0.053
	(0.032)	(0.040)	(0.054)
N	26,480	16,576	9904
Limiting condition			
Treatment effect	−0.034	−0.027	−0.028
	(0.020)	(0.023)	(0.035)
N	23,170	14,504	8666
False treatment			
Treatment effect	0.033	0.044	0.021
	(0.032)	(0.039)	(0.054)
N	24,825	15,540	9285

Standard errors in brackets. Estimated on balanced panels.
** $p < 0.05$.

Table 3
Treatment effect by gender.

	Model 1 All	Model 2 Never smoked = 1	Model 3 Never smoked = 0
Women			
Treatment effect	0.014	**0.121****	−0.107
	(0.047)	(0.061)	(0.074)
N	12,192	7856	4336
Men			
Treatment effect	0.047	0.074	−0.004
	(0.044)	(0.052)	(0.079)
N	14,288	8720	5568

Standard errors in brackets. Estimated on balanced panels.
** $p < 0.05$.

They find health benefits, but only for non-smoking women, suggesting the importance of a reduction in exposure to second-hand smoke.