Income Convergence and Regional Labor Markets in the U.S. Revisited

Soohyun Choi

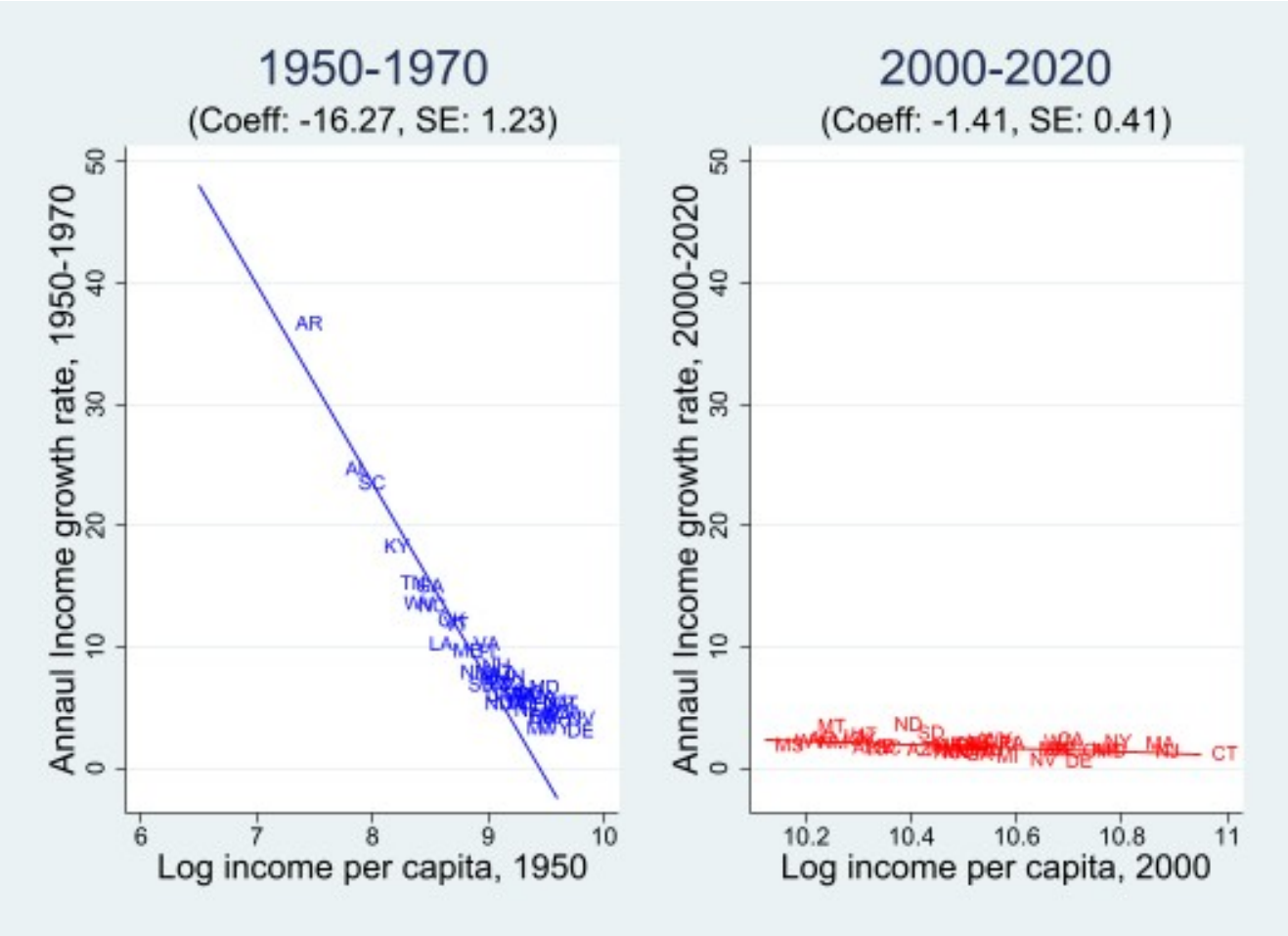
University of Texas at Dallas

December 19, 2022

Contact: Soohyun.Choi@utdallas.edu

Motivation

Classic beta convergence test on U.S. state income: Barro & Sala-i-Martin (2003), Ganong & Shoag (2017)



Motivation

- According to the classic beta convergence test, the rate of U.S. income convergence across states has declined for the last three decades, unlike the strong convergence in the mid-20th century
- Problem: Beta convergence is a necessary condition of convergence, not a sufficient condition (Sul, 2019)

$\rightarrow$ Revisit the question by applying a relative convergence test

Data: Income per capita

- Personal income per capita from the Bureau of Economic Analysis
- Calendar years 1947-2020
- 48 states in the contiguous United States
- Units are 2020 dollars after being deflated by CPI

Relative convergence test

Phillips and Sul (2007): If the slope coefficients on the time trend are identical in the long run, the growth rate of the log per capital real income converges to the steady-state level

- Suppose X_{it} is the panel data of state i at time t

$$X_{it} = \delta_{it} u_t \tag{1}$$

where δ_{it} is a time-varying idiosyncratic element and u_t is a single common component.

- Removing the common factor as follows:

$$h_{it} = \frac{X_{it}}{\frac{1}{N}\sum_{i=1}^{N} X_{it}} = \frac{\delta_{it}}{\frac{1}{N}\sum_{i=1}^{N} \delta_{it}} \tag{2}$$

where h_{it} is the relative transition parameter, tracing out a transition path of state i in relation to the panel average

Relative convergence test

- Convergence condition is that the variance of h_{it} converges to zero.

$$H_{it} = \frac{1}{N}\sum_{i=1}^{N}(h_{it} - 1)^2 \to 0 \text{ if } \lim_{t\to\infty} \delta_{it} = \delta \text{ for all } i \tag{3}$$

- $\lim_{t\to\infty} \delta_{it} = \delta$ for all i in eq (3) is equivalent to

$$\lim_{t\to\infty} \frac{X_{it}}{X_{jt}} = 1 \text{ for all } i \text{ and } j \tag{4}$$

Relative convergence test

- The hypothesis test can be implemented through the following log t regression model (Null hypothesis: Convergence)

$$log\frac{H_1}{H_t} - 2log(log(t)) = a + b\ log(t) + \epsilon_t \tag{5}$$

$$\text{for } t = [rT], [rT]+1, ..., T \text{ with } r > 0$$

- Recommend to set r=1/3 according to Phillips & Sul (2007)
- However, setting r=1/10 gives a more conservative result (Kwak, 2022)
- The null of convergence cannot be rejected if $t_{\hat{b}} \geq -1.65$ at the 5% level (one side test)

Relative convergence test: Result

- About 30% of early period samples are discarded to eliminate the impact of initial observations
- The log t results can vary for different measurement units $\rightarrow$ Subtract the minimum value across states and time to solve the measurement unit problem
- Log t regression is performed after eliminating the trend and cyclical components of a times series of each state

Table: Log t regression

	Income per capita (1971-2020)	
	Phillips & Sul (2007)	Kwak (2022)
$\hat{b}$	-0.175	-0.520
$t_{\hat{b}}$	-3.604	-135.093
Convergence	No	No

Convergence clubs

Even though we reject the null hypothesis of convergence for the whole panel, convergent subgroups - *clubs* - can still exist

Step 1: Cross-section sorting

Sort states in descending order according to the incomes in the last period. Index states with their orders $\{1, ..., N\}$

Step 2: Core group formation

- Find the first k such that $t_k > -1.65$ for the subgroup with states $\{k, k+1\}$. In case of no k satisfying $t_k > -1.65$, there are no subgroups in the panel
- If the first convergence subgroup exists, perform log t regression for the subgroups with states $\{k, k+1, ..., k+j\}$ with $j \in \{1,...,N\text{-}k\}$. A group of states $\{k, k+1, ..., k+j^*\}$ yielding the highest value of the test statistics form a core group.

Convergence clubs

Step 3: Sieve states for club membership

- Form a complementary group $G_{j^*}^c$ within which states are not in the core group. Add one state from $G_{j^*}^c$ at a time to the core group and run log t test. If the test statistic is greater than the critical value $c^* \in [-1.65, 0]$, include the state in the club candidate group.
- Run the log t test for the club candidate group above. If the test statistic is greater than -1.65, the initial convergence club is obtained. If not, raise the critical value c^* and repeat the step 3 until $\hat{t_b} > -1.65$. If there are some $\hat{t_b} > -1.65$, add one state from the remaining candidates at a time, run the log t. If the highest value of $\hat{t_b}$ is not greater than -1.65, stop the procedure. The extended core group forms an initial convergence club. Otherwise, repeat the above procedure to add the state with the highest $\hat{t_b}$.

Convergence clubs

Step 4: Recursion and stopping rule

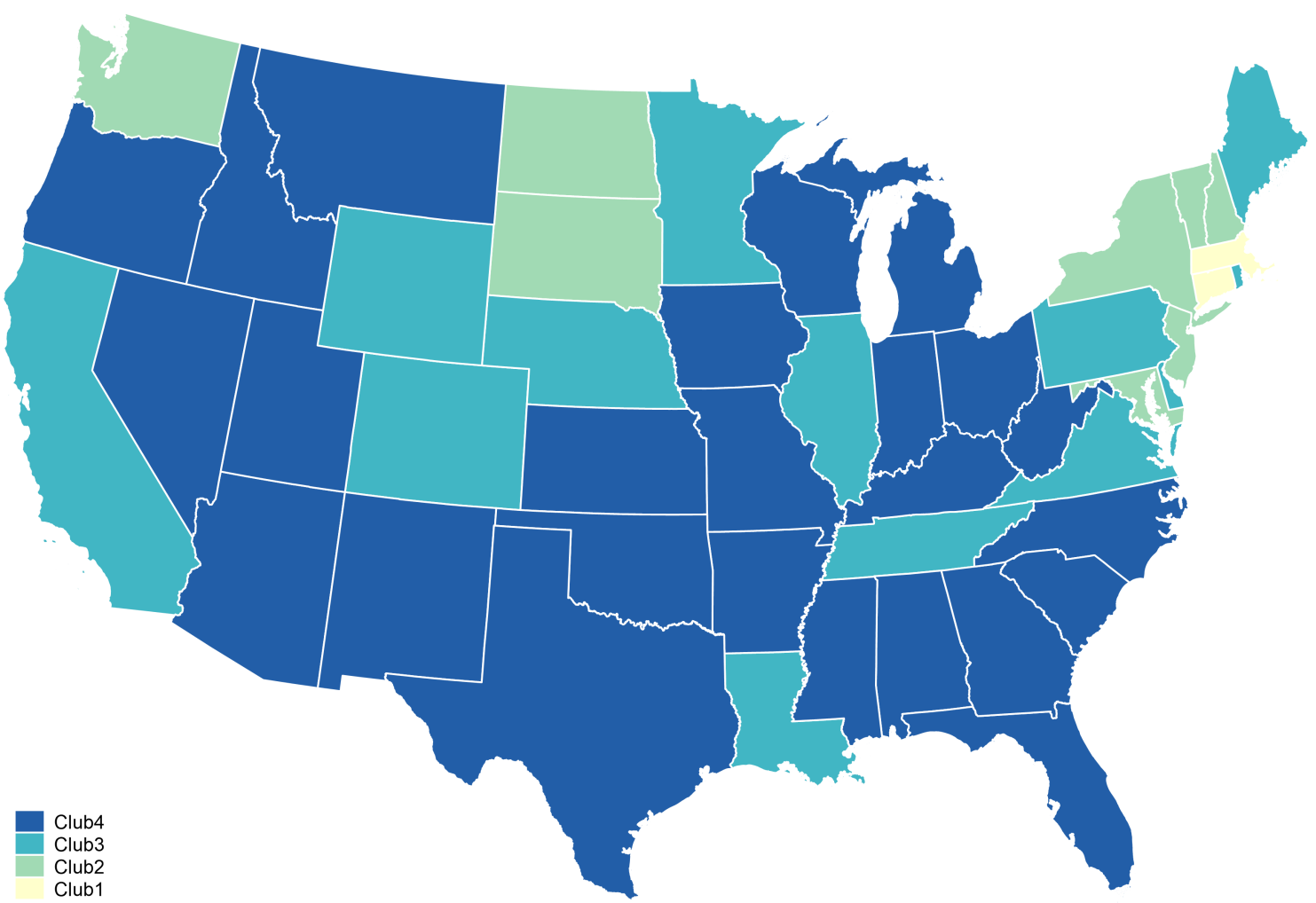
Perform the log t test for the remaining states that are not sieved by step 3. If the test statistic is greater than -1.65, the subgroup becomes another convergence club. Otherwise, repeat steps 1-3 on this subgroup.

Step 5: Club merging

Run the log t test for all pairs of the subsequent initial clubs. Merge clubs satisfying the convergence hypothesis jointly. Repeat the procedure on the new clubs until no clubs can be merged.

Convergence clubs: Results

Figure: Geographic Distribution of Convergence Clubs: Income per Capita



Possible determinants of income divergence

Focusing on labor mobility

1. Housing supply (Ganong & Shoag, 2017): Land use regulation
2. Changes in labor demand from skill-biased technological change (Autor & Dorn, 2013): Percentage employment of high-skilled jobs
3. Local amenities, such as public education spending, crime rate, air quality, weather (Diamond, 2016; Rappaport, 2007): Public K-12 spending per student, Violent crime incidents per 100 Inhabitants, Air quality index, Daylight hours

Possible determinants: Data Appendix

Variable	Sources	Year	Note
Percentage employment of high-skilled jobs	U.S. Census ACS 5-year estimate	2010-2021	- High-skilled jobs are equivalent to cognitive nonroutine occupations - The occupations include: Management, Business and financial operation Computer and engineering Life/physical/social science Community and social service Legal Education/training/library Arts/design/entertainment/sports/media Healthcare practitioners and technical (Dvorkin & Shell, 2017)
Violent Crimes	FBI Uniform Crime Reports	1985-2020	- Violent crime incidents per 100 Inhabitants - Include homicide, rape, robbery, and aggravated assault
Daylight Hours	National Centers for Environment Information	Cross-Sectional	The total time that sunshine reaches the surface of Earth is expressed as the percentage of the maximum amount possible from sunrise to sunset with clear sky conditions
K-12 spending	U.S. Census	2002-2020	- Public K-12 spending per student
Land Use Regulation	Gyourko et al. (2008, 2021)	Cross-Sectional	- Wharton Residential Land Use Regulatory Index - Use the time series average of the two data sets
Air Quality	U.S. Environmental Protection Agency	1980-2022	- Takes into account all of the criteria air pollutants measured (C.O., NO_2, O_3, PM2.5, PM10)

Possible determinants: Summary statistics

Table: Summary Statistics of Determinants

	Convergence Clubs			
	Club 1	Club 2	Club 3	Club 4
Per Capita Income (2020 Dollars)	45,675	36,948	34,256	28,991
	(1,427)	(584)	(376)	(216)
Land Use Regulation Index	0.61	0.2	-0.01	-0.3
	(0.04)	(0.03)	(0.02)	(0.01)
High-Skilled Employment (%)	43.32	39.59	37.16	34.39
	(0.44)	(0.34)	(0.26)	(0.14)
Average Sunny Days (%)	58	51.33	59.88	61.59
	(0.00)	(0.3)	(0.28)	(0.23)
Violent Crimes per 100 Inhabitants	0.43	0.35	0.43	0.46
	(0.02)	(0.01)	(0.01)	(0.01)
K-12 Spending per Student (Thousand Dollars)	14.84	13.23	11.26	8.95
	(0.53)	(0.35)	(0.17)	(0.08)
Air Quality Index	41.63	33.86	37.4	37.15
	(0.56)	(0.37)	(0.31)	(0.21)
Number of Member States	2	8	13	25

Note: Standard errors are reported in parentheses.

Cross-sectional and random effect regression

Table: Impacts of Housing supply, Labor Demand Change and Amenities on U.S. Regional Income

Model	Log Income Per Capita					
	Cross-sectional	Cross-sectional	RE with year dummy	RE with year dummy	RE with year dummy	RE with year dummy
	(1)	(2)	(3)	(4)	(5)	(6)
Land Use Regulation	0.137*	0.122***	0.078***	0.024	0.022	0.022
	(0.035)	(0.034)	(0.029)	(0.026)	(0.026)	(0.026)
Daylight Hours (%)		-0.341**	-0.045	0.085	0.064	0.066
		(0.157)	(0.134)	(0.140)	(0.140)	(0.142)
K-12 Spending per student ($1,000)			0.305***	0.320**	0.312**	0.309**
			(0.080)	(0.062)	(0.065)	(0.064)
High-Skilled Employment (%)				0.582**	0.601**	0.595**
				(0.217)	(0.220)	(0.226)
Air Quality Index					0.069*	0.070*
					(0.039)	(0.040)
Violent Crimes per 100 inhabitants						-0.006
						(0.026)
Observations	48	46	874	506	506	506
R^2	0.254	0.320	0.639	0.748	0.746	0.745

A problem of random effect regression

Some variables are stationary, while others are non-stationary
$\rightarrow$ Run ordered logit where a convergence club classification is a dependent variable

Table: Appendix:
Unit Root Test for Panel Data Using Augmented Dickey-Fuller Tests (Choi, 2001)

Variables	Inverse χ^2	Inverse Normal	Inverse Logit	Modified Inverse χ^2	Note
Income Per Capita (2020 Dollars)	378.9265***	-10.9833***	-14.0150***	20.4185***	At least one panel is stationary
High-Skilled Employment (%)	181.3113***	0.3305	-1.2684	6.1568***	
Violent Crimes per 100 Inhabitants	108.9401	-1.0018	-1.0639	0.9339	All panels contain unit roots
K-12 Spending per Student (Thousand Dollars)	118.5337*	0.7047	0.5784	1.6262*	
Air Quality Index	285.7648***	-6.8531***	-9.1743***	13.6951***	At least one panel is stationary

Ordered logit regression

Figure: Ordered Logit Estimation Results: Coefficients

	Final Clubs					
	(1)	(2)	(3)	(4)	(5)	(6)
Land Use Regulation Index	1.450***	1.255**	0.446	-0.629	-0.646	-0.639
	(0.551)	(0.568)	(0.636)	(0.765)	(0.812)	(0.818)
Daylight Hours (%)		-5.107**	1.602	1.239	1.152	1.127
		(2.214)	(2.880)	(3.008)	(3.354)	(3.374)
K-12 Spending per student ($1000)			8.622***	7.372***	7.330***	7.375***
			(2.220)	(2.398)	(2.495)	(2.604)
High-Skilled Employment (%)				19.183***	19.310***	19.323***
				(6.729)	(7.078)	(7.082)
Air Quality Index					0.173	0.149
					(2.934)	(2.958)
Violent Crimes per 100 Inhabitants						0.060
						(0.985)
Observations	48	46	46	46	46	46
Log-Likelihood	-49.893	-45.561	-34.581	-29.506	-29.504	-29.502
χ^2(d.f.)	8.173	12.885	34.844	44.995	44.999	45.002
Prob > χ^2	0.004	0.002	0.000	0.000	0.000	0.000
Pseudo R^2	0.076	0.124	0.335	0.433	0.433	0.433

Ordered logit regression

Figure: Ordered Logit Estimation Results: Odd Ratios

	Final Clubs					
	(1)	(2)	(3)	(4)	(5)	(6)
Land Use Regulation Index	4.265	3.509	1.562	0.533	0.524	0.528
Daylight Hours (%)		0.006	4.964	3.453	3.164	3.087
K-12 Spending per Student ($1000)			5554	1590	1525	1595
High-Skilled Employment (%)				∞	∞	∞
Air Quality Index					1.189	1.161
Violent Crimes per 100 Inhabitants						1.062

Conclusion

- Log t test fails to reject the hypothesis of U.S. income convergence across states
- Alternatively, income convergence has formed among four subgroups – called *clubs* – within which states have specific characteristics in common.
- The ordered logit regression analysis suggests that changes in labor demand and public school spending play a significant role in accounting for the formation and composition of convergence clubs.