



UNIVERSITY OF WARSAW

Faculty of Economic Sciences

Is the Gini Coefficient Enough?

A Microeconomic Data Decomposition Study

Ivan Skliarov, PhD student

Inequality literature

1. Correlation between income inequality and education, health, economic growth;
2. As a control variable;
3. Inequality decomposition studies: gender, age, race, regions, income sources.

Human development index



Summary measure of average achievement in key dimensions of human development:

1. Healthy life;
2. Knowledge;
3. Decent standard of living.

Source: hdr.undp.org

Human development index



A **composite index** measuring average achievement in key dimensions of human development:

1. Life expectancy at birth;
2. Mean years of schooling;
3. Expected years of schooling;
4. GNI per capita.

Source: hdr.undp.org

Human development index



Mean years of schooling – average N of years of education received by people ages 25 and older.

Expected years of schooling – N of years of schooling that a child can expect to receive.

Source: hdr.undp.org

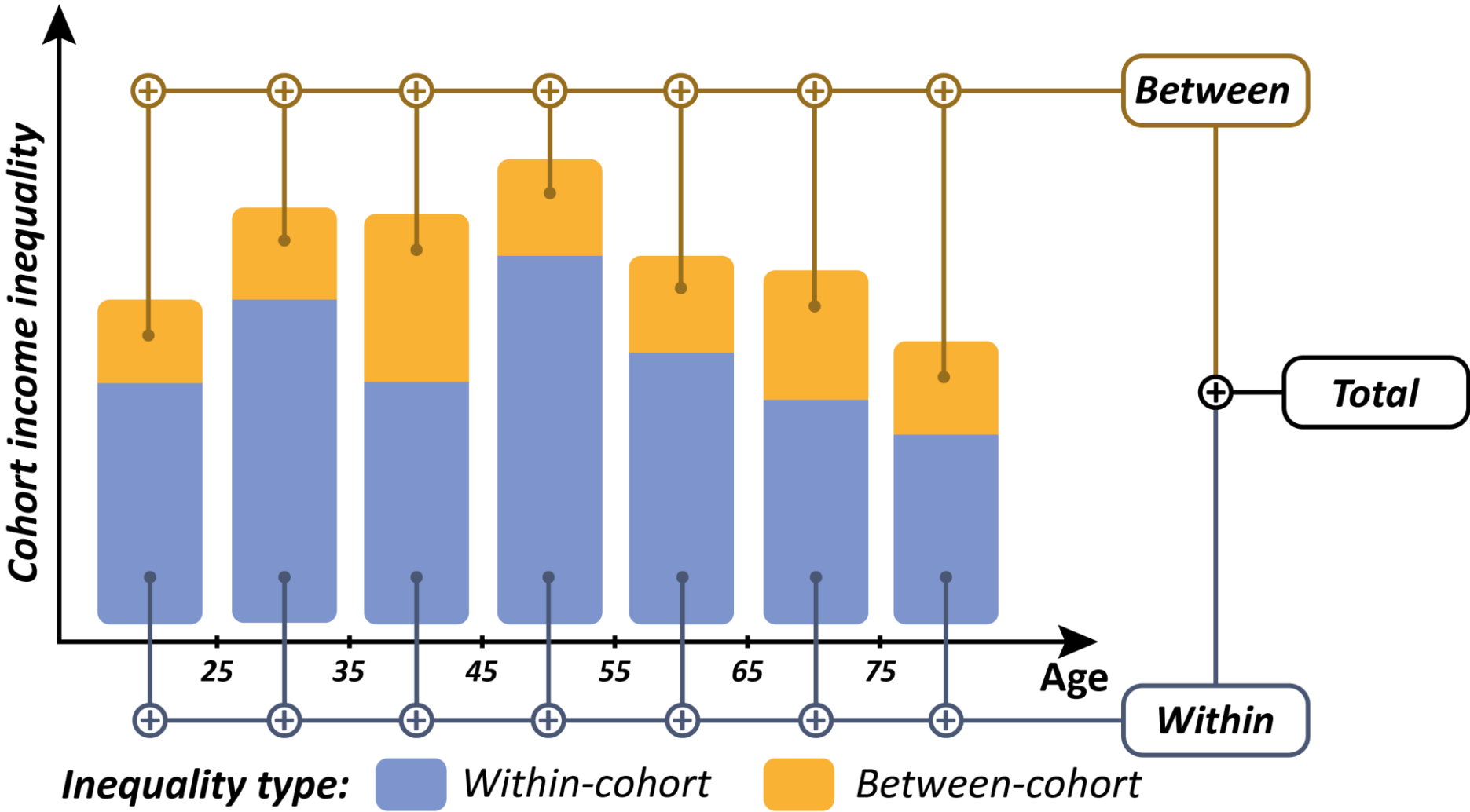
Methodology

- 5 dependent variables:
 - **Human Development Index;**
 - Life expectancy;
 - Mean years of schooling;
 - Expected years of schooling;
 - GDP p.c.
- 2 indexes: Thiel or squared coefficient of variation
- 2 sets of control variables: Castells-Quintana et al. (2018) and Amate-Fortes et al. (2017).
- 2 estimators: OLS with PCSE and GLS.

Income data

- Disposable household income;
- Top and bottom-coded;
- Equivalized;
- Time period: 2006-2017;
- 18 countries: Austria, Belgium, Canada, Colombia, France, Germany, Ireland, Israel, Luxembourg, Peru, Poland, Spain, Switzerland, the UK, the US, Uruguay, Georgia, Lithuania.

Inequality indexes composition



Source: Own

Theil Index

$$Theil = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{\mu} \ln \left(\frac{y_i}{\mu} \right)$$

Decomposition:

$$Theil = \sum_{k=1}^K \frac{n_k}{n} \frac{\mu_k}{\mu} \left[\frac{1}{n_k} \sum_{i=1}^{n_k} \frac{y_{i,k}}{\mu_k} \ln \left(\frac{y_{i,k}}{\mu_k} \right) \right] + \sum_{k=1}^K \frac{n_k}{n} \frac{\mu_k}{\mu} \ln \left(\frac{\mu_k}{\mu} \right)$$

Theil Index

Decomposition:

The size of k -th cohort

Equalized disposable income

$$Theil = \sum_{k=1}^K \frac{n_k}{n} \frac{\mu_k}{\mu} \left[\frac{1}{n_k} \sum_{i=1}^{n_k} \frac{y_{i,k}}{\mu_k} \ln \left(\frac{y_{i,k}}{\mu_k} \right) \right] + \sum_{k=1}^K \frac{n_k}{n} \frac{\mu_k}{\mu} \ln \left(\frac{\mu_k}{\mu} \right)$$

The total population size

Average income of the k -th cohort

Average income of the entire population

The diagram illustrates the decomposition of the Theil Index. The formula is presented as $Theil = \sum_{k=1}^K \frac{n_k}{n} \frac{\mu_k}{\mu} \left[\frac{1}{n_k} \sum_{i=1}^{n_k} \frac{y_{i,k}}{\mu_k} \ln \left(\frac{y_{i,k}}{\mu_k} \right) \right] + \sum_{k=1}^K \frac{n_k}{n} \frac{\mu_k}{\mu} \ln \left(\frac{\mu_k}{\mu} \right)$. Four text labels with arrows point to specific parts of the formula: 'The size of k -th cohort' points to n_k in the first sum; 'Equalized disposable income' points to $\frac{y_{i,k}}{\mu_k}$ in the inner sum; 'The total population size' points to n in the denominator of the first sum; 'Average income of the k -th cohort' points to μ_k in the denominator of the inner sum; 'Average income of the entire population' points to μ in the denominator of the second sum.

Theil Index

Decomposition:

Individual contributions of k -th cohort members
to within-cohort inequality

$$Theil = \sum_{k=1}^K \underbrace{\left[\frac{n_k}{n} \frac{\mu_k}{\mu} \right]}_{\text{Weight}} \left[\frac{1}{n_k} \sum_{i=1}^{n_k} \frac{y_{i,k}}{\mu_k} \ln \left(\frac{y_{i,k}}{\mu_k} \right) \right] + \sum_{k=1}^K \underbrace{\left[\frac{n_k}{n} \frac{\mu_k}{\mu} \right]}_{\text{Weight}} \underbrace{\ln \left(\frac{\mu_k}{\mu} \right)}_{\text{Contributions of } k\text{-th cohort to between-cohort inequality}}$$

Squared coefficient of variation

$$SCV = \frac{\sigma^2}{2\mu^2}$$

Decomposition:

$$SCV = \sum_{k=1}^K \frac{n_k}{n} \left(\frac{\mu_k}{\mu}\right)^2 \frac{\sigma_k^2}{2\mu_k^2} + \frac{1}{2} \sum_{k=1}^K \frac{n_k}{n} \left[\left(\frac{\mu_k}{\mu}\right)^2 - 1 \right].$$

Squared coefficient of variation

Decomposition:

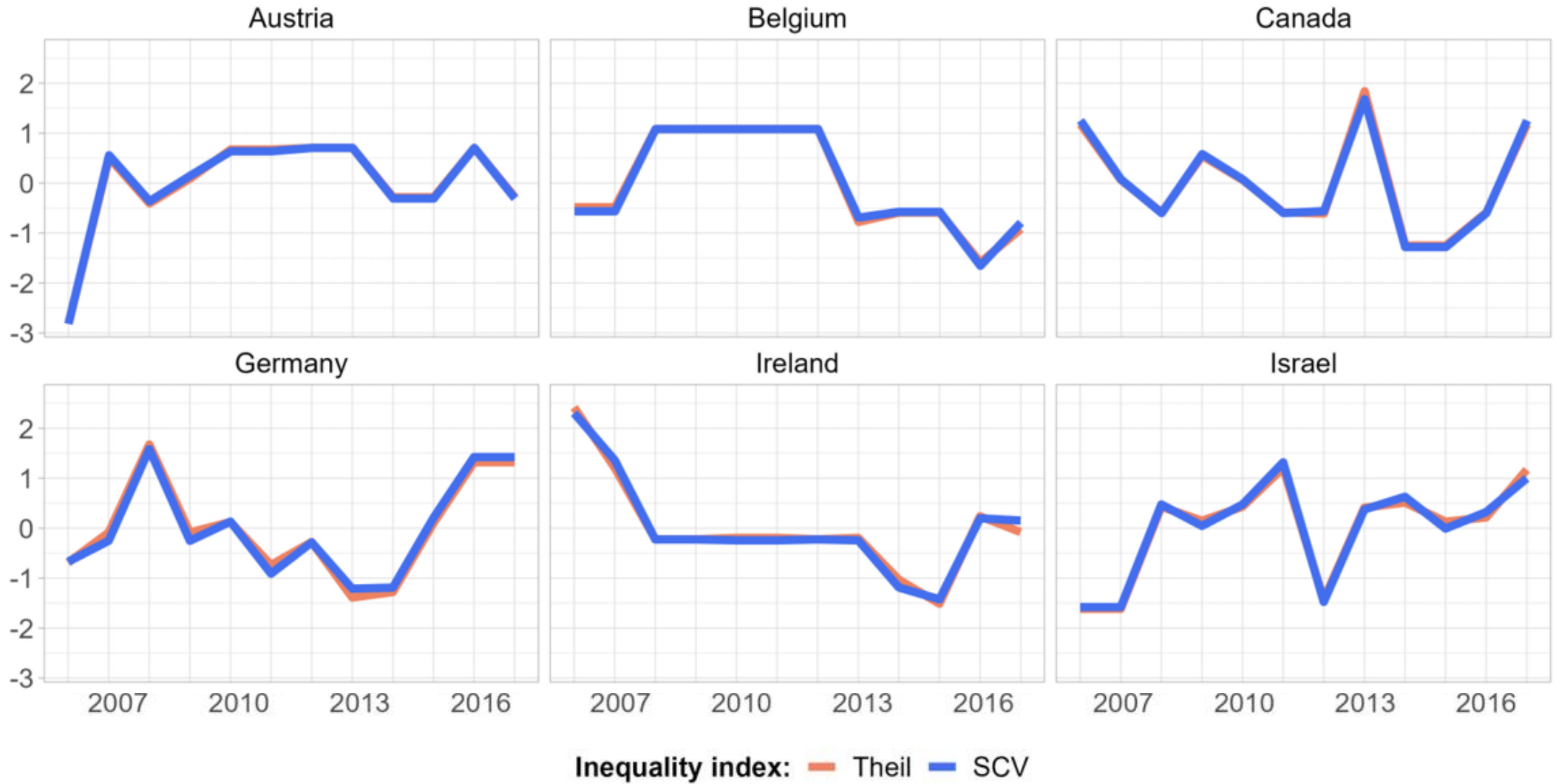
Contribution of k -th cohort to within-cohort inequality

$$SCV = \sum_{k=1}^K \left[\frac{n_k}{n} \left(\frac{\mu_k}{\mu} \right)^2 \right] \left[\frac{\sigma_k^2}{2\mu_k^2} \right] + \frac{1}{2} \sum_{k=1}^K \left[\frac{n_k}{n} \right] \left[\left(\frac{\mu_k}{\mu} \right)^2 - 1 \right].$$

Weight

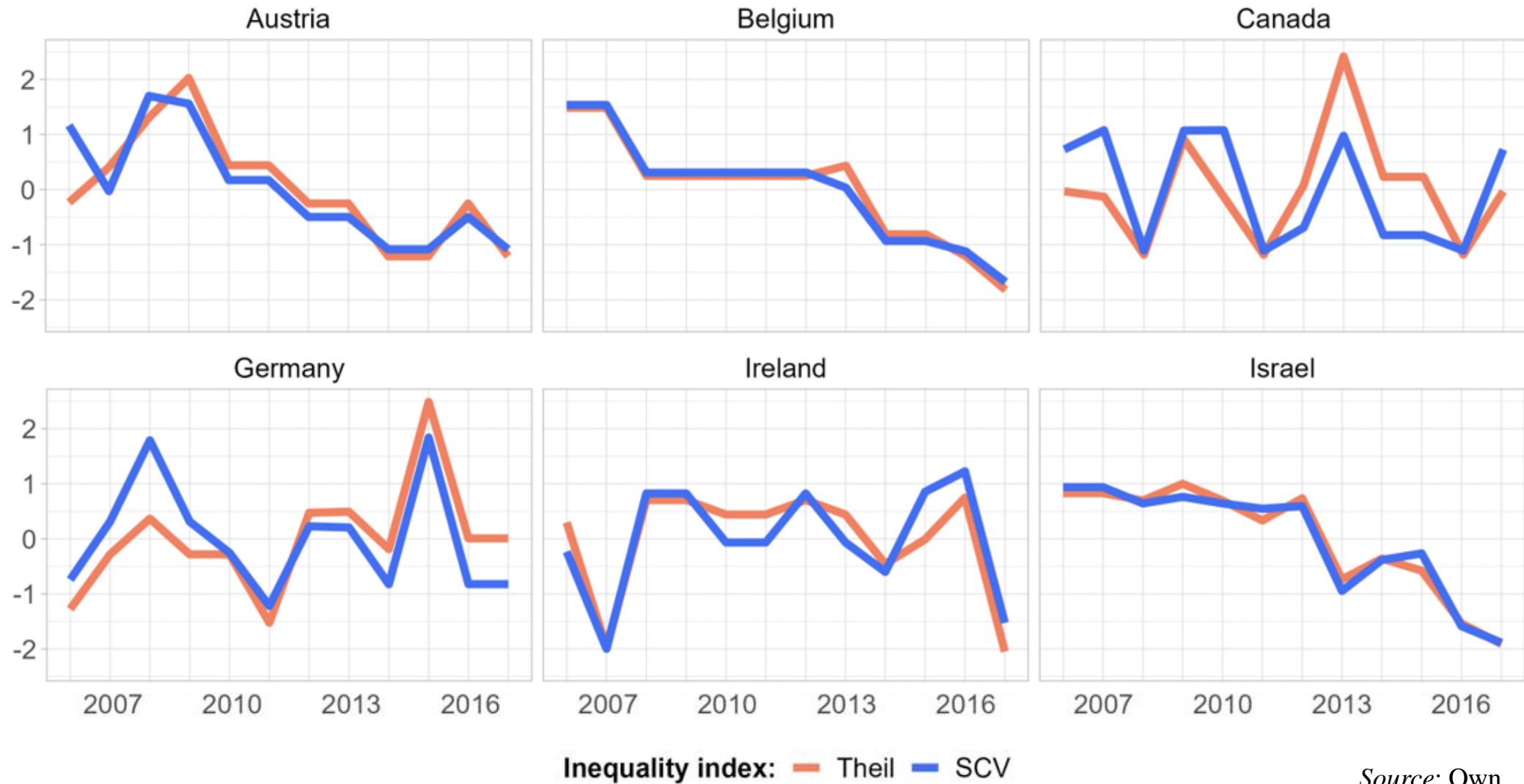
Contribution of k -th cohort to between-cohort inequality

Comparison of between-cohort indexes



Source: Own

Comparison of within-cohort indexes



Source: Own

List of control variables

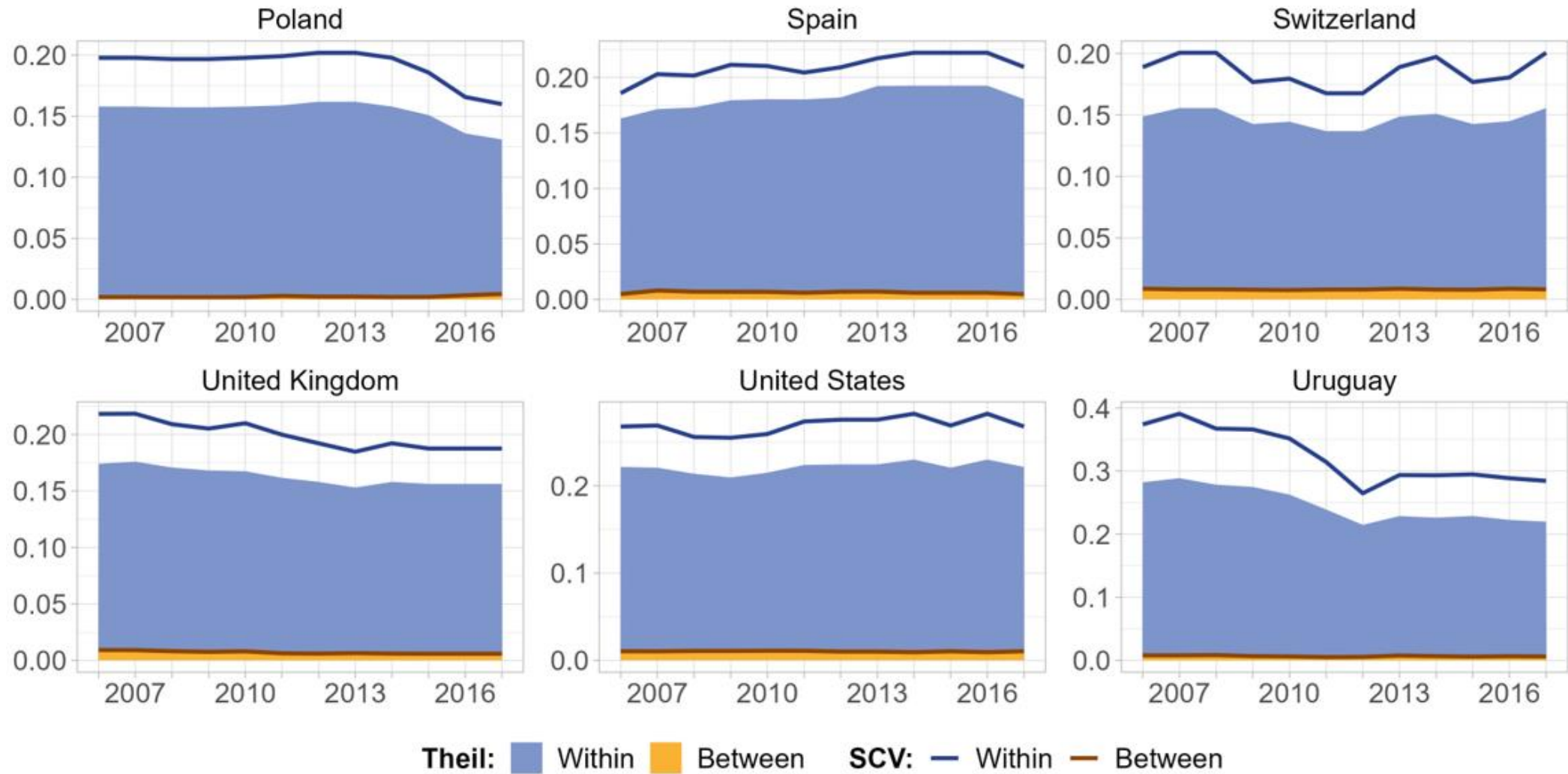
Model A – Castells-Quintana et al. (2018)

Model B – Amate-Fortes et al. (2017)

Description	Models	Source
Inflation, GDP deflator (annual %)	A	WDI
Gross capital formation (% of GDP)	A	WDI
Sum of imports and exports of goods and services (% of GDP)	A	WDI
Urban population (% of total population)	A	WDI
General government final consumption expenditure (% of GDP)	A, B	UNCTAD
Population growth (annual %)	B	WDI
Unemployment, total (% of total labor force)	B	WDI
Control of Corruption: Estimate	B	WGI
Political Stability and Absence of Violence/Terrorism: Estimate	B	WGI

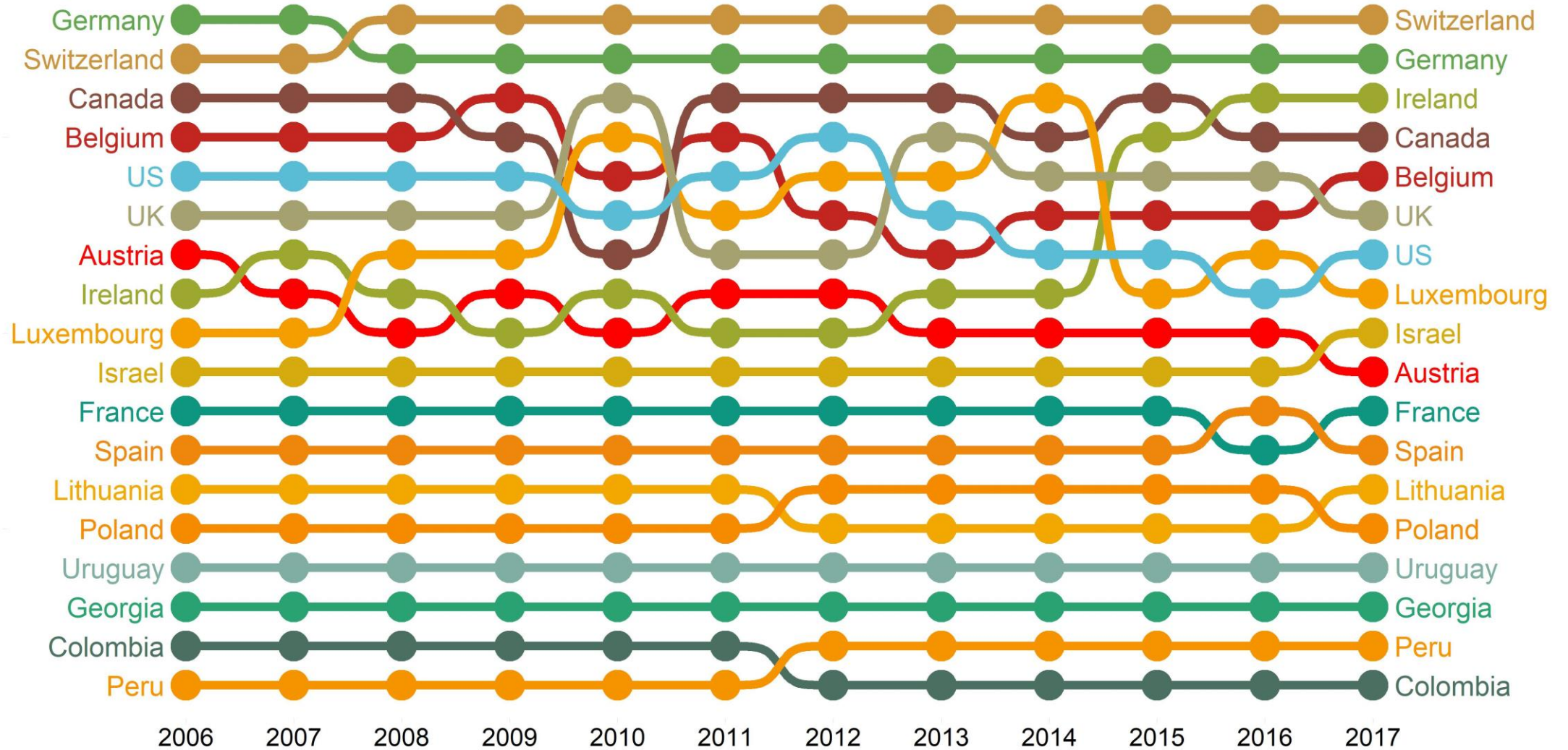
Source: Own

Within and between-cohort inequality



Source: Own

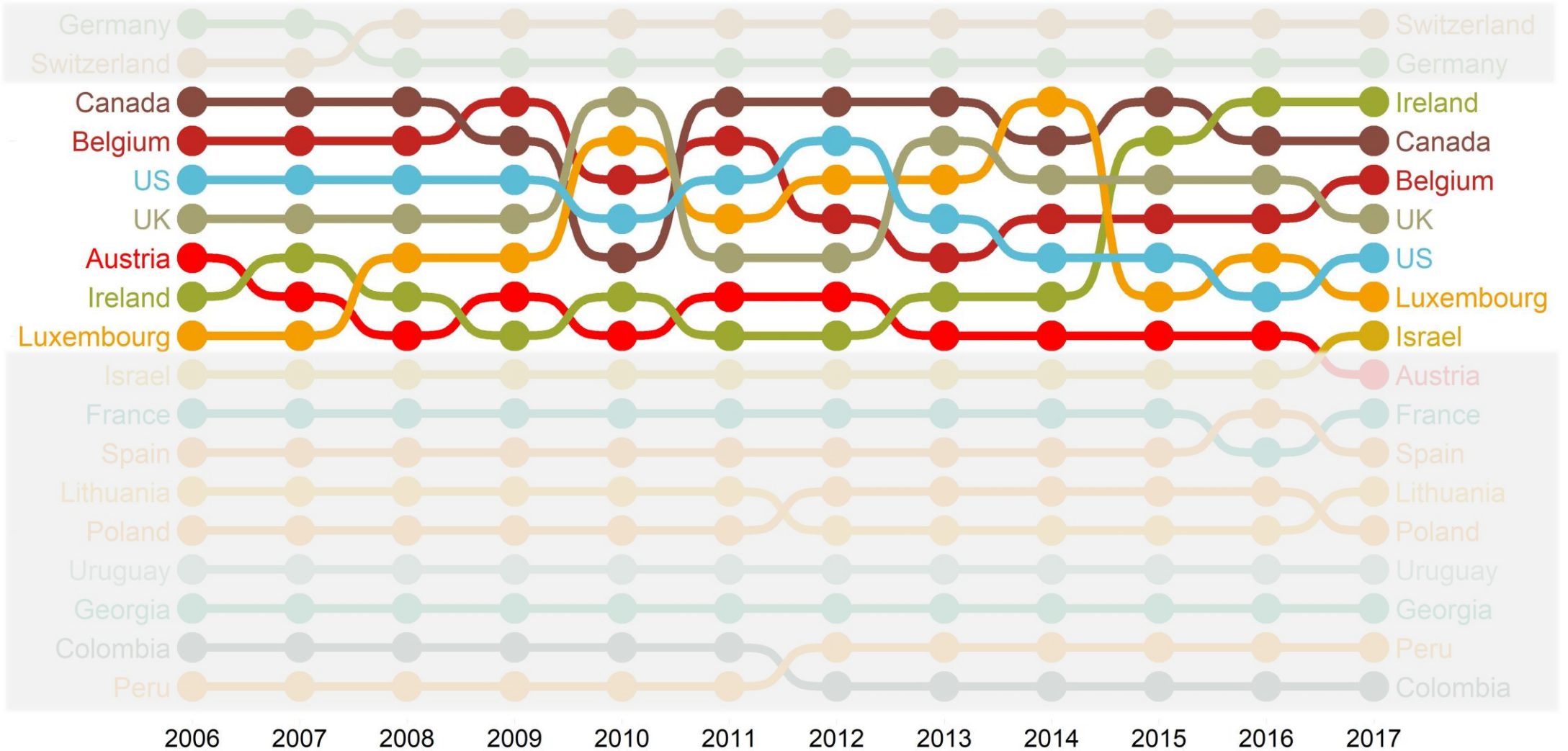
HDI Ranking



Note: Countries ranked from the highest to the lowest HDI

Source: Own

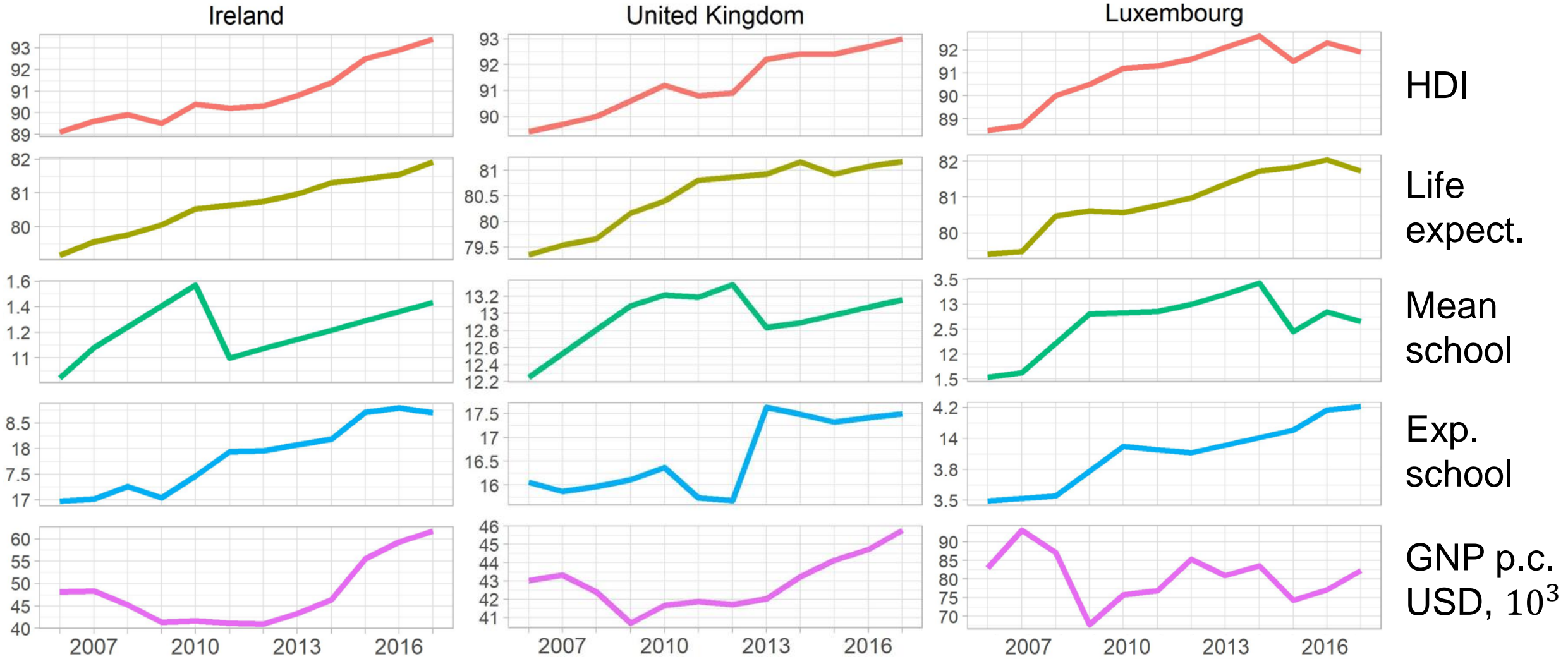
HDI Ranking



Note: Countries ranked from the highest to the lowest HDI

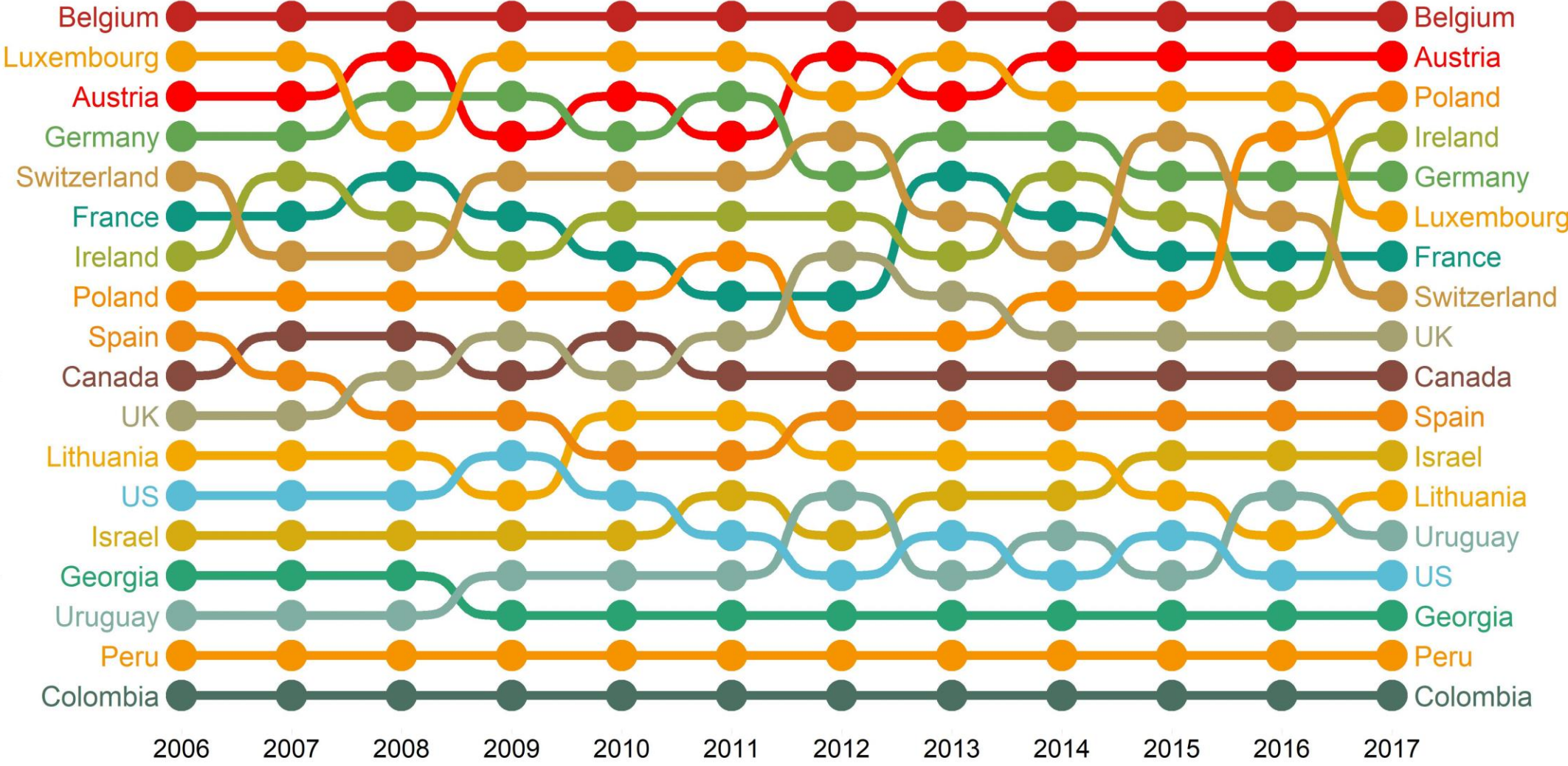
Source: Own

Dynamic of HDI and its components, 2006-2017



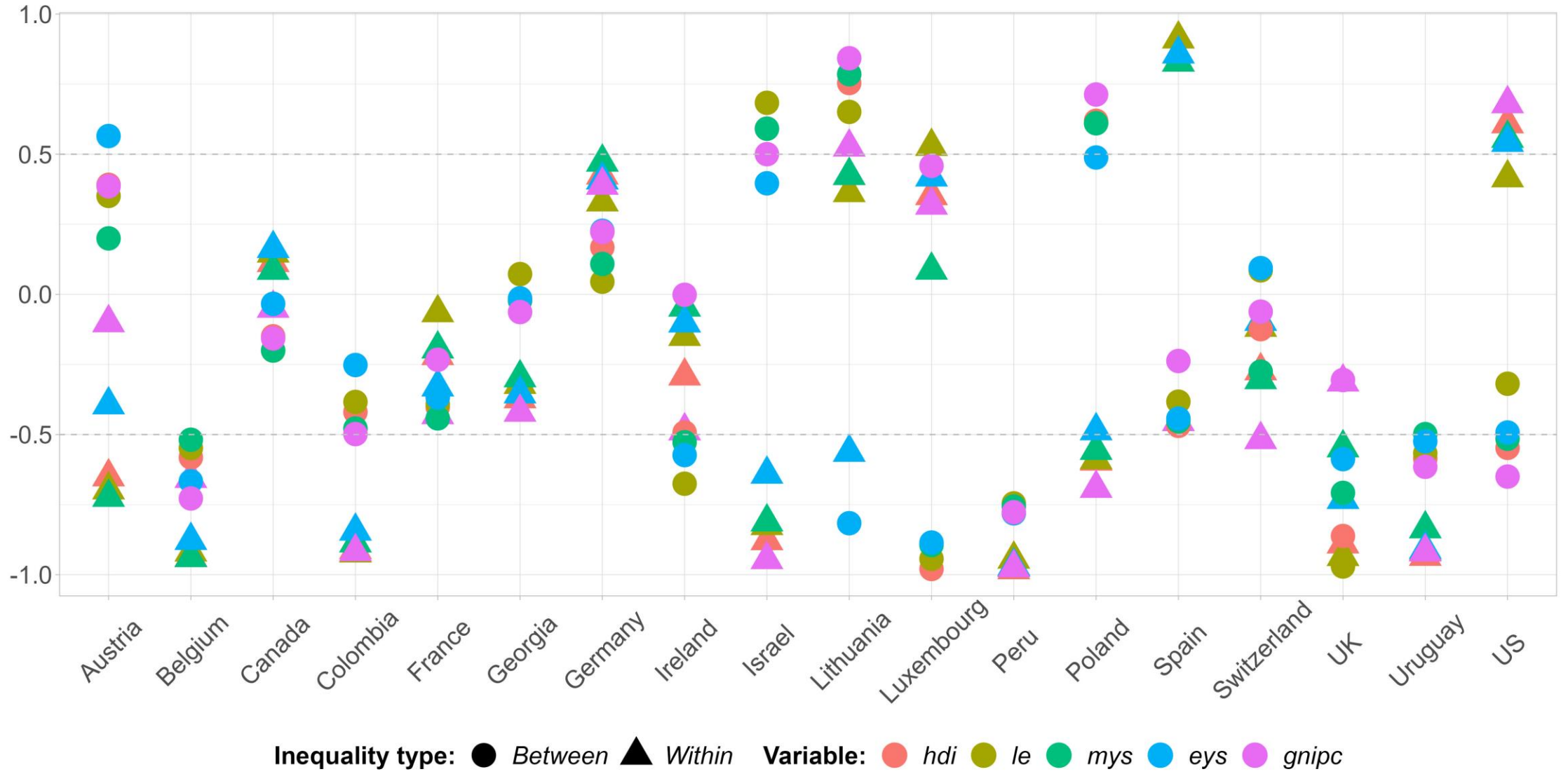
Source: Own

Income inequality ranking, Theil index



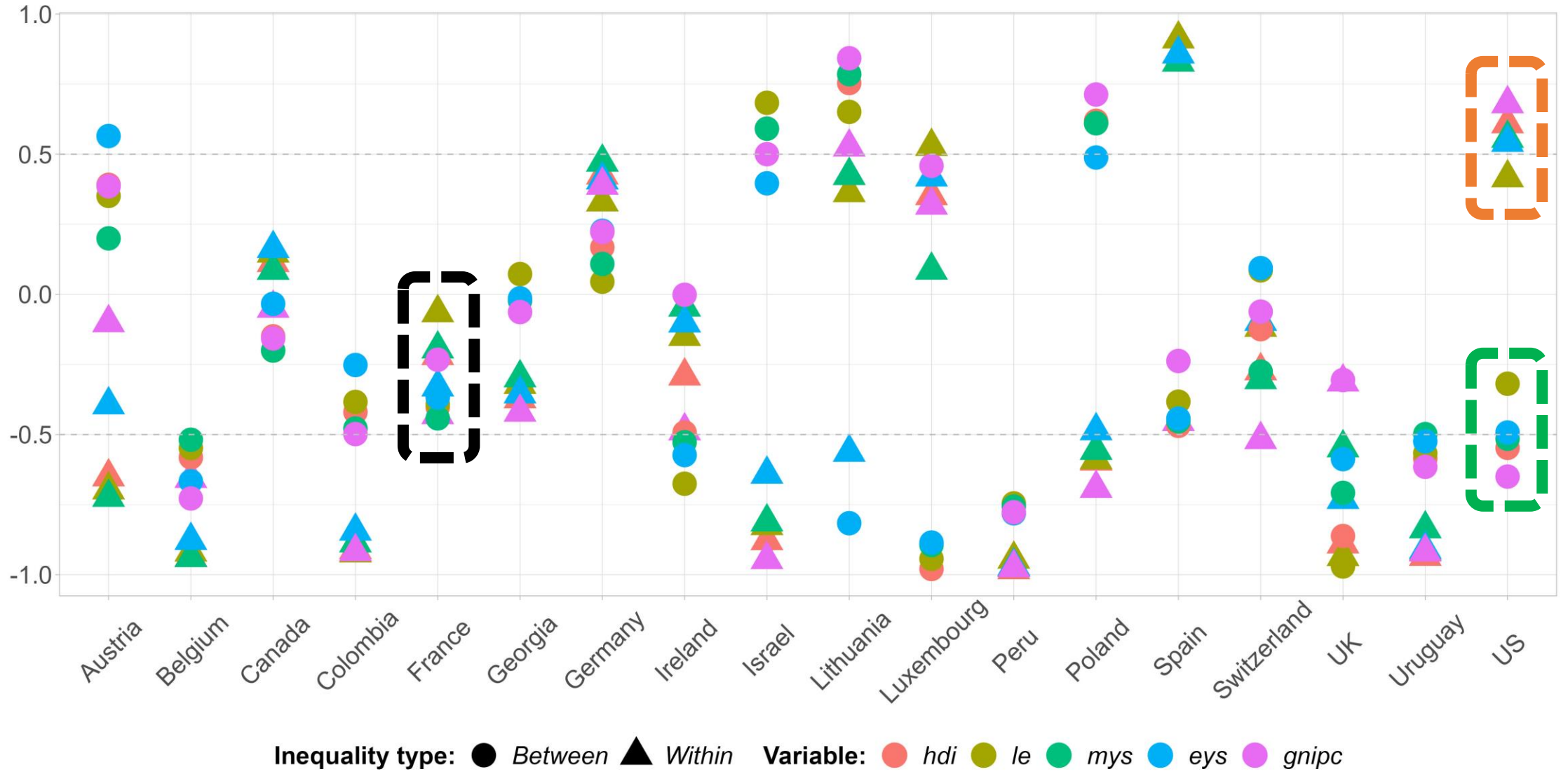
Note: Countries ranked from the most equal to the most unequal.
 Source: Own

Correlation of inequality components with HDI



Source: Own

Correlation of inequality components with HDI



Source: Own

Estimation results, Theil index, OLS with PCSE

x	Model A					Model B				
	<i>hdi</i>	<i>le</i>	<i>mys</i>	<i>eys</i>	<i>gnipc</i> ¹	<i>hdi</i>	<i>le</i>	<i>mys</i>	<i>eys</i>	<i>gnipc</i> ¹
W_t	-13.9***	-5.1***	-2.6*	-1.4	20.6	-11.4***	-2.2	-3.5***	-0.6	30.6
W_{t-1}	-10.3***	-3.2*	1.2	-4.2**	-69.6**	-8.9***	-3.0*	0.9	-4.1**	-39.6*
W_{t-2}	-4.9	-3.4	-0.1	-1.6	23.8	-4.6	-2.6	0.3	-1.4	-9.0
B_t	16.8	8.8	-1.6	-21.6	195.1	-1.9	17.1	0.4	-32.1**	213.9*
B_{t-1}	38.0	38.6**	3.4	-33.8**	44.7	9.3	39.3***	10.0	-46.8***	-0.9
B_{t-2}	27.1	22.3	4.5	-23.2*	234.1*	-5.9	6.0	10.0	-29.4**	-123.3
d	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
g	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	18	18	18	18	18	18	18	18	18	18
T	12	12	12	12	12	12	12	12	12	12

Coefficients for intercept, control, and indicator variables are omitted.

Statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

¹ Coefficients measured in thousands.

Source: Own calculation using LIS data.

Comparison of regression results obtained using OLS with PCSE

x	Theil					SCV				
	<i>hdi</i>	<i>le</i>	<i>mys</i>	<i>ey</i> s	<i>gnipc</i>	<i>hdi</i>	<i>le</i>	<i>mys</i>	<i>ey</i> s	<i>gnipc</i>
W_t	-A -B	-A	-a -B			-A -B	-A	-a -B		
W_{t-1}	-A -B	-a -b		-A -B	-A -b	-A -B	-A		-A -B	-A
W_{t-2}										
B_t				-B	+b				-B	+a +b
B_{t-1}		+A +B		-A -B			+A +B		-A -B	
B_{t-2}				-a -B	+a				-a -B	

Statistical significance: capital letter means $p < 0.05$, small letter means $p < 0.1$.

Source: Own calculation using LIS data.

Comparison of regression results obtained using the GLS estimator

x	Theil					SCV				
	<i>hdi</i>	<i>le</i>	<i>mys</i>	<i>eys</i>	<i>gnipc</i>	<i>hdi</i>	<i>le</i>	<i>mys</i>	<i>eys</i>	<i>gnipc</i>
W_t	-A -B	-A	-B			-A -B	-A	-B	-a	-a
W_{t-1}	-A -B	-A -b		-A -B		-A -B	-A		-A -B	
W_{t-2}						-b	-b		-b	
B_t				-A -B	+b				-A -B	+B
B_{t-1}	+A	+A +B		-A -B	+A	+a	+A +b	+b	-A -B	+A
B_{t-2}	+A				+A	+A			+B	+a -b

Statistical significance: capital letter means $p < 0.05$, small letter means $p < 0.1$.

Source: Own calculation using LIS data.

Extensions

- More countries;
- More inequality dimensions.

ineq.2d: Two-Dimensional Decomposition of the Theil Index and the Squared Coefficient of Variation

<https://cran.r-project.org/web/packages/ineq.2d/index.html>

Theil index - Giammatteo, M. (2007)

Squared coefficient of variation - Garcia-Penalosa, C., & Orgiazzi, E. (2013)

Ineq.2d

```
theil.2d(us16, "hitotal", "cohort", c("hilabour", "hicapital", "hitransfer"), "hpopwgt")
```

Source	f25t49.W	f50t74.W	f75.W	t24.W	f25t49.B	f50t74.B	f75.B	t24.B
hilabour	0.14081	0.16409	0.00268	0.02948	0.0401	0.01272	-0.00516	-0.00498
hicapital	0.00684	0.02248	0.00188	-0.00009	0.00118	0.00174	-0.0022	-0.00002
hitransfer	-0.01202	-0.01897	0.00564	-0.00139	0.00258	0.00422	-0.0291	-0.00069

The rest of R and STATA code is publicly available:

<https://doi.org/10.6084/m9.figshare.c.6665180.v1>

References

- Amate-Fortes, I., Guarnido-Rueda, A., & Molina-Morales, A. (2017). Economic and Social Determinants of Human Development: A New Perspective. *Social Indicators Research*, 133(2), 561-577.
- Castells-Quintana, D., Royuela, V., & Thiel, F. (2019). Inequality and Sustainable Development: Insights from an Analysis of the Human Development Index. *Sustainable Development*, 27(3), 448-460.
- Garcia-Penalosa, C., and Orgiazzi, E. (2013). Factor Components of Inequality: A Cross-Country Study. *Review of Income and Wealth*, 59(4), 689-727.
- Giammatteo, M. (2007). The Bidimensional Decomposition of Inequality: A nested Theil Approach. LIS Working papers, Article 466, 1-30.
- Mookherjee, D., & Shorrocks, A. (1982). A Decomposition Analysis of the Trend in UK Income Inequality. *The Economic Journal*, 92(368), 886-902.

Thank you!