Appendix to APSR Article: Symbolic Refugee Protection: Explaining Latin America's Liberal Refugee Laws

Omar Hammoud-Gallego *and Luisa Feline Freier †

19/05/2022

1 Introduction

This annex to the paper "Symbolic Refugee Protection: Explaining Latin America's Liberal Refugee Laws" is subdivided as follows: in section 2 we present the summary statistics and the distribution of our dependent variable. In section 3, we apply a series of statistical tests to scrutinise the structure of our data, in order to justify our choice of regression models, both non-spatial and spatial.

In the section 4, we then report the results from ordinary least square (OLS) panel data model. We also show results from OLS panel data models with the dependent variable lagged by one and three years. In section 5 and 6, we show results from Poisson and Quasipoisson models, also using the same specifications as in the Tobit models presented in the main article. The former are used in count data models, whereas the latter are used as they take into account the overdispersion of our data. We report the results from all these different models to show how even using slightly different methods, the direction – and often the statistical significance – of the explanatory variables in our models do not change. Last, in the section 7 we show results from our linear spatial panel data models, disaggregating between direct and indirect effects, and clarify the difference between our SAR (Spatial Autoregressive) and SEM (Spatial Error) models.

2 Model Variables

In this section we begin by presenting the summary statistics of the variables included in our models (Table 1).

Some issues arise: first, – as shown in Figure 1 – our dependent variable is not normally distributed, and a simple log is not useful due to the presence of too many zero (given that most

*Fellow in Political Science and Public Policy, School of Public Policy, The London School of Economics and Political Science, o.hammoud-gallego@lse.ac.uk, orcid-id: 0000-0001-9723-1451

†Associate Professor of Social and Political Sciences, Department of Social and Political Science, Universidad del Pacifico (Lima, Peru), lf.freierd@up.edu.pe, orcid-id: 0000-0002-4653-4812

Table 1: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Regulatory Complexity	551	40.30	28.15	0	82
VDEM Polyarchy	551	0.65	0.20	0.08	0.93
Left-Wing Gov	551	0.40	0.49	0	1
Change in GDP Per Capita	551	1.93	3.73	−18.17	16.26
Trade as % of GDP	551	61.08	29.43	14	167
International Migration Stock	551	2.24	2.40	0.00	13.40
Refugees as % of pop.	551	0.13	0.55	0.00	8.85
Emigrants in US and Spain	551	3.77	4.04	0.06	21.53

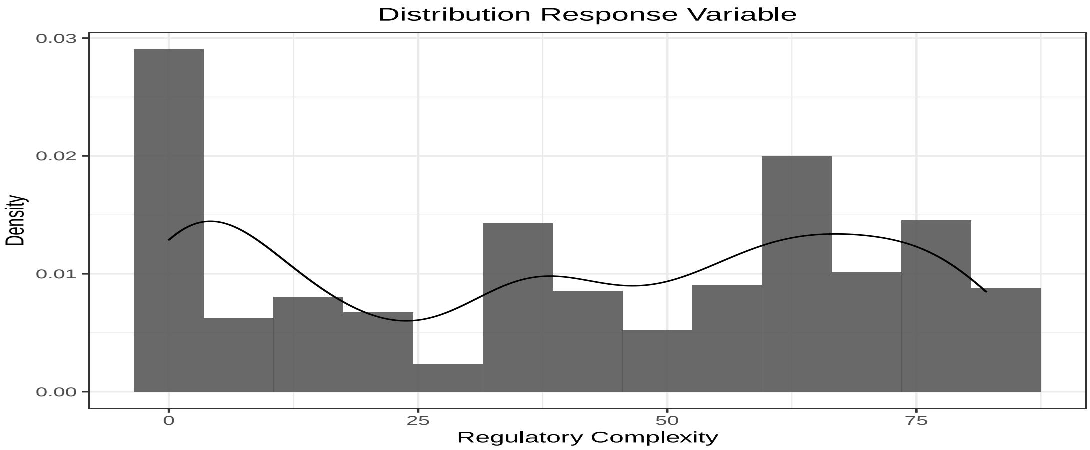

Figure 1: Regulatory Complexity Distribution. Source: APLA Database.

countries in Latin America did not have meaningful asylum legislation until the early 2000s). It is for this reason that we decided to apply a Tobit model as the main model, and then use a Poisson panel data regression model here in the appendix in which we interpret our data as count data. Similarly, in our explanatory variables (not shown here), we see that while some variables have somewhat normally distributed data, most of them have either long tails, or are heavily skewed towards the left – in most cases. Also, in most cases, the presence of zeros in various variables (e.g. for the variable refugee as a proportion of the population in the host country) suggests that logging them is not a recommendable option. Nonetheless, in the Poisson models, we can log (log_e) our dependent variable

$$\eta_{i,t} = log(\mu_{i,t})$$

as that allows us to assume that the transformed mean of our dependent variable follows a linear model.

3 Structure of the Data and Best Model Fit

In this section we perform a series of statistical tests to understand the structure of our data and which models fit best. we test contemporaneous correlation of the residuals across the countries included in our model through the application of two tests: the Breusch-Pagan LM test of independence and the Pesaran CD test (Hsiao, Pesaran, and Pick 2012). Given our results, we reject the null hypotheses of no cross-sectional error correlations in both tests, thus confirming that there is indeed cross-sectional dependence among the countries considered in our sample (respectively, $\chi^2 = 906.2150674$, df = 171, p-value= $1.8699109 \times 10^{-100}$; z=-3.5280754, p-value= 4.1859289×10^{-4}). We also conduct a Breusch-Godfrey/Wooldridge test for serial correlation – which might lead our results to have smaller standard errors and higher R^2 coefficients than they are ($\chi^2 = 327.3143932$, df = 29, p-value= $3.0645852 \times 10^{-52}$). We apply it and reject the null hypothesis of no serial correlation, which we must then consider in our results. Moreover, to check for stochastic trends, we conduct a Dickey-Fuller test, which confirms that our series is stationary (Dickey-Fuller= -6.2194549, Lag = 2, p-value = 0.01). Finally, after conducting a Breusch-Pagan test, we reject the null hypothesis of homoskedasticity, and therefore confirm the detection of heteroskedasticity (Breusch-Pagan= 125.8112289, df= 53, p-value = 7.4043262×10^{-8}). Given the results of the tests above, in our Tobit models the standard errors are clustered at the country level, whereas in the OLS (in Appendix) we apply robust estimators (HC4) (Millo 2017).

In addition to the above, we conduct a Hausman test to check whether to use random effect or fixed effects models ($\chi^2 = 127.8851378$, df = 7, p-value= $1.7379474 \times 10^{-24}$). The test checks whether the unique errors are correlated with the regressors, and the null hypothesis is that the unique errors do not correlate with the regressors (Dougherty 2016). As the p-value is significant, we use fixed effects rather than random effects as our models of choice. Additionally, we apply the F test for individual effects and the Breusch-Pagan test for balanced panels to check the need for time-fixed effects (respectively, F=13.896684, df1 and df2= 28, 497, p-value = $3.0112036 \times 10^{-46}$; $\chi^2 = 920.979129$, df = 2, p-value= $1.0278344 \times 10^{-200}$) (Croissant and Millo 2008). We reject the null hypothesis that no time-fixed effects are needed and therefore include them in our fixed-effects regressions. Finally, we run a Monte Carlo simulation of Moran I test – a measure of spatial autocorrelation – to test the relation between the values of our dependent variable and the location where it is measured. To calculate Moran's I, we first build an inverse matrix of the distance between the different countries of the region based on their coordinates, and then run 1000 simulations of the test (Statistic=0.0084067, Observed Rank=997, p-value=0.003). Based on its results we reject the null hypothesis of zero spatial correlation in our dependent variable and therefore complement our earlier models with a series of spatial panel data models to test

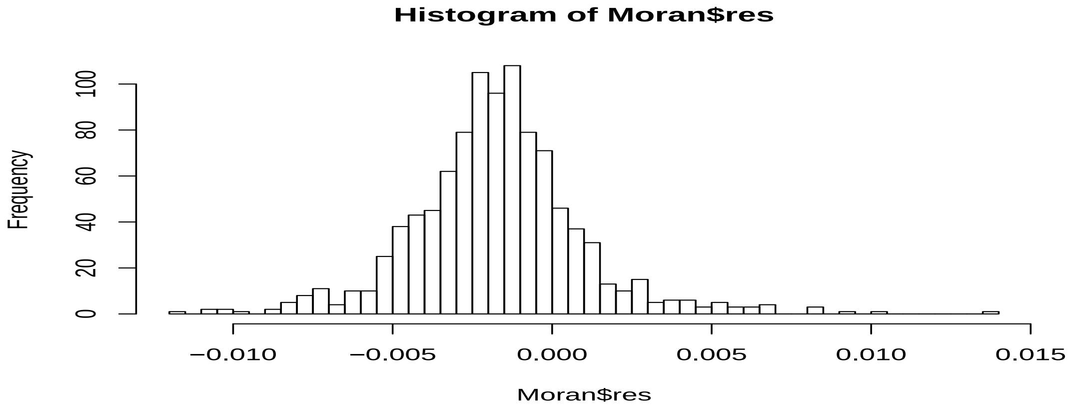

Figure 2: Moran I Residuals

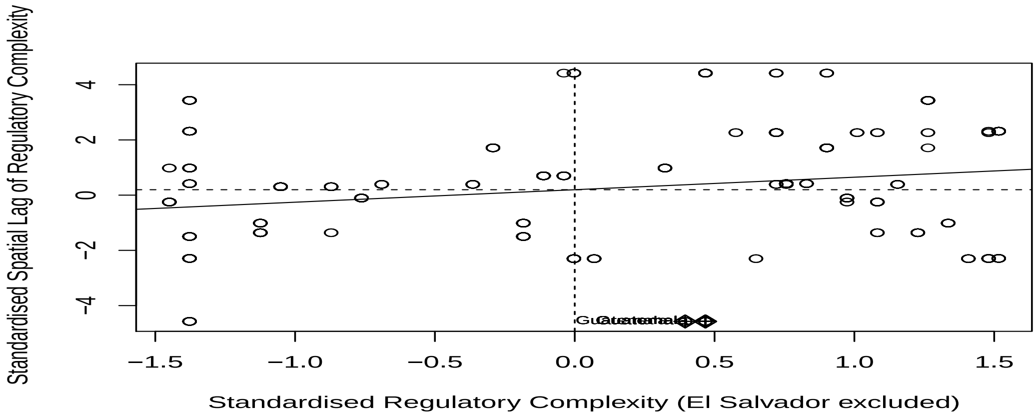

Figure 3: Plot Moran I Test. El Salvador is excluded as it is a statistical outlier.

the determinants of regulatory complexity and liberalisation, while taking into account spatial spill overs (LeSage and Pace 2009; Ward and Gleditsch 2008).

Figure 2 shows the distribution of the Moran I residuals, whereas Figure 3 plots the relation between our dependent variable Regulatory Complexity on the x-axis, and the same variable but spatially lagged on the y-axis. The slope of fit equals Moran I. The upper-right and lower-left quadrants represent positive spatial correlation, that is, countries that are geographical neighbours have similar values. Opposite to these, the upper-left and lower-right quadrants represent negative spatial correlation, whereby countries close to each other have dissimilar values. In both axes the variable is standardised. We exclude El Salvador from the plot, as it seems to be a big outlier, and its inclusion does not allow to perceive the actual pattern in the data.

4 Linear Models

In this section we present the results of our OLS panel data models following the specification of the Tobit models specified in the paper. We present these results to show that regardless of the model used, our coefficients' directions remain as expected. Table 2 and Table 3 report the results of standard OLS panel data models, without and with one and three year lagged responsive variable respectively. We do not standardise the coefficients of our regression models to avoid "apples to oranges" comparisons (King 1986). We apply HC4 estimators, as suggested by Cribari-Neto (2004) and Zeileis (2004). We do not apply the function vcovHAC as it cannot be used for panel data (Millo 2017).

Table 2: OLS Panel Data Regression

	Dependent variable: Regulatory Complexity				
	(1)	(2)	(3)	(4)	(5)
VDEM Polyarchy	-7.85				-3.99
	(26.94)				(26.89)
Left-Wing Gov	17.42***				15.84***
	(4.56)				(4.42)
Change in GDP Per Capita		-0.14			-0.09
		(0.26)			(0.25)
Trade as % of GDP		0.38**			0.31**
		(0.16)			(0.13)
International Migration Stock			-0.82		-0.49
			(2.35)		(1.99)
Refugees as % of pop.				0.70	1.40
				(2.79)	(1.67)
Emigrants in US and Spain					-0.62
					(1.82)
Fixed-effects					
Country	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Year	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	551	551	551	551	551
R^2	0.14	0.07	0.001	0.0004	0.18
Adjusted R^2	0.06	-0.02	-0.09	-0.09	0.09
Note:Robust SE in parentheses			*p<0.1; **p<0.05; ***p<0.01		

Table 3: OLS Panel Data Regression (Lag 1 and 3 years)

	Dependent variable: Regulatory Complexity			
	One Year Lag		Three Year Lag	
	(1)	(2)	(3)	(4)
VDEM Polyarchy		−1.45		−3.89
		(28.55)		(28.63)
Left-Wing Gov	14.80***	14.70***	13.14***	12.97***
	(4.05)	(3.87)	(2.93)	(3.18)
Change in GDP Per Capita		−0.21		0.26
		(0.25)		(0.40)
Trade as % of GDP	0.27**	0.28**	0.21*	0.22*
	(0.12)	(0.13)	(0.12)	(0.13)
International Migration Stock		−0.72		−1.40
		(1.72)		(2.23)
Refugees as % of pop.		1.81		0.11
		(4.87)		(18.04)
Emigrants in US and Spain		−0.64		−0.55
		(1.92)		(2.23)
Country Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Year Fixed Effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	532	532	494	494
R^2	0.15	0.15	0.11	0.12
Adjusted R^2	0.06	0.06	0.02	0.02

Note:Robust SE in parentheses *p<0.1; **p<0.05; ***p<0.01

5 Poisson Models

To further confirm our findings, we apply a series of Poisson models used with count data. Poisson regression are usually applied to account for the non-normal distribution of the dependent variable, and assume its variance to be a function of the mean. However, as our dependent variable is over-dispersed, below we also use a series of Quasipoisson models to further confirm our findings.

Table 4: Poisson Panel Data Regression on Regulatory Complexity with Country-Year Fixed Effects. Source: APLA Database, V-Dem Database, Political Institutions Database, UN DESA, World Bank, authors' own estimates.

Dependent Variable:	Regulatory Complexity				
Coefficients:	Incidence Rate Ratio				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
VDEM Polyarchy	0.76				0.89
	(0.72)				(0.81)
Left-Wing Gov	1.56***				1.47***
	(0.24)				(0.21)
Change in GDP Per Capita		1.00			1.00
		(0.01)			(0.01)
Trade as % of GDP		1.01**			1.01*
		(0.01)			(0.01)
International Migration Stock			0.85		0.86
			(0.18)		(0.13)
Refugees as % of pop.				0.95	1.08
				(0.23)	(0.12)
Emigrants in US and Spain					0.97
					(0.05)
Fixed-effects					
Country	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	551	551	551	551	551
Squared Correlation	0.661	0.646	0.649	0.638	0.676
Pseudo R^2	0.51345	0.50386	0.49336	0.48457	0.53321
BIC	8,631.22	8,789.19	8,949.39	9,094.14	8,368.86

One-way (Country) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

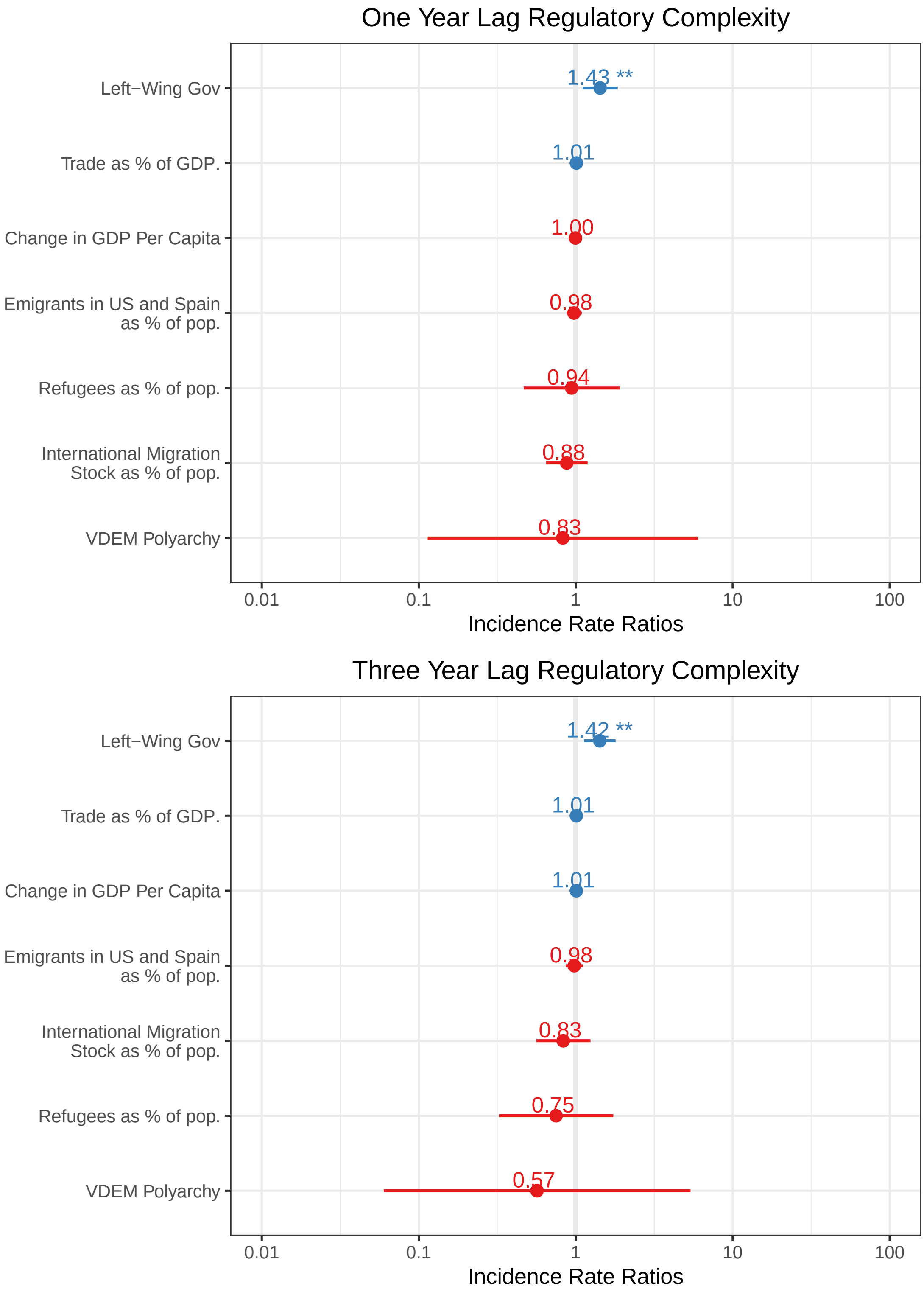

Figure 4: Source: APLA Database, V-Dem Database, Political Institutions Database, UN DESA, World Bank, authors' own estimates.

6 Quasipoisson Models

In this section we present results from Quasipoisson models. These models are used to account for the overdispersion present in our data. Overdispersion is found when the main assumption of Poisson models, i.e. that $mean(Y) = \sigma^2(Y)$ is not met. In our case, the Quasipoisson model reported in Table 5, model 5, shows that the dispersion parameter is around 8.9, that is, σ^2 is eight time the mean. This confirms the presence of overdispersion. Table 6 reports the results of Quasipoisson regressions with the dependent variable lagged by one and three years.

Table 5: Quasipoisson Regression

	Dependent variable: Regulatory Complexity				
	(1)	(2)	(3)	(4)	(5)
VDEM Polyarchy	−0.28				−0.12
	(0.30)				(0.44)
Left-Wing Gov	0.45***				0.38***
	(0.07)				(0.06)
Change in GDP Per Capita		−0.0004			0.0002
		(0.01)			(0.01)
Trade as % of GDP		0.01***			0.01***
		(0.002)			(0.002)
International Migration Stock			−0.16**		−0.15
			(0.07)		(0.09)
Refugees as % of pop.				−0.05	0.08
				(0.34)	(0.47)
Emigrants in US and Spain					−0.03
					(0.02)
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	551	551	551	551	551
Dispersion Parameter	9.26	9.36	10.7	10.01	8.92
Note: Robust SE in parentheses			*p<0.1; **p<0.05; ***p<0.01		

Table 6: Quasipoisson Regression (Lag 1 and 3 Years)

	Dependent variable: Regulatory Complexity			
	One Year Lag			Three Year Lag
	(1)	(2)	(3)	(4)
VDEM Polyarchy		−0.19		−0.57
		(0.36)		(0.37)
Left-Wing Gov	0.37***	0.36***	0.37***	0.35***
	(0.07)	(0.07)	(0.07)	(0.07)
Change in GDP Per Capita		−0.003		0.01
		(0.01)		(0.01)
Trade as % of GDP	0.01***	0.01***	0.01***	0.01***
	(0.002)	(0.002)	(0.002)	(0.002)
International Migration Stock		−0.13**		−0.18**
		(0.07)		(0.09)
Refugees as % of pop.		−0.06		−0.29*
		(0.23)		(0.17)
Emigrants in US and Spain		−0.02		−0.02
		(0.02)		(0.02)
Country Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	532	532	532	494

Note: Robust SE in parentheses *p<0.1; **p<0.05; ***p<0.01

7 Linear Spatial Panel Data Models

In this last section we briefly clarify the difference between our Spatial Autoregressive Models (SAR) and the Spatial Error Models (SEM). Additionally, we present our SAR models – both with fixed effects and random effects – disaggregating the effects by type: direct, or indirect. We start with our SEM model that is specified as follows:

$$Y_{i,t} = \beta_0 + \gamma_{i,t} + \delta_{i,t} + \zeta_{i,t} + \eta_{i,t} + \theta_{i,t} + \kappa_{i,t} + \nu_{i,t} + \alpha_i + \xi_t + \upsilon_{i,t}$$

$$\upsilon_{i,t} = \lambda W \upsilon_{i,t} + \epsilon_{i,t}$$

The model resembles a standard OLS model, except that υ includes the weight matrix W and spatial coefficient λ that measures the average strength of spatial correlation among the error terms. On the other hand, the SAR model implies that the changes in an explanatory variable in any geographical point will affect the value of the dependent variable regardless of the location of the latter. Further discussions on the characteristics of these models are beyond the scope of this paper's research question. Last, we clarify why we show our SAR models disaggregated by effect type in Tables 7 to 9. As Golgher and Voss (2016) explain, direct effects represent the "the expected average change across all observations for the dependent variable in a particular region due to an increase of one unit for a specific explanatory variable in that region." Opposite to this concept are indirect effects, which represent "changes in the dependent variable of a particular region arising from a one unit increase in an explanatory variable in another region." In our case, what they define as "regions" are the countries in Latin America considered in our study. Therefore, indirect effects show how changes in some of the explanatory variables in one country effectively spill over into another.

Table 7: Regulatory Complexity Spatial Panel Data Models. Main Effects.

	(1) SAR FE	(2) SAR RE
Main		
V-DEM Polyarchy	-2.43	-1.27
	(7.58)	(7.42)
Left-Wing Gov	15.5***	18.3***
	(1.86)	(1.87)
Change in GDP per Capita	-0.100	-0.14
	(0.20)	(0.19)
Trade as % of GDP	0.29***	0.26***
	(0.059)	(0.054)
International Migration Stock as % of Population	-0.71	0.055
	(1.03)	(0.87)
Refugees as % of Population	1.33	0.29
	(1.56)	(1.54)
Emigrants in US and Spain as % of Pop.	-0.69	0.44
	(0.57)	(0.47)
Spatial		
ρ	-0.38**	0.67***
	(0.14)	(0.034)
Variance		
σ^2_e	201.0***	239.7***
	(12.2)	(14.9)
lgtθ		-1.37***
		(0.22)
Observations	551	551
R^2	0.088	0.347

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Regulatory Complexity Spatial Panel Data Models. Direct and Indirect Effects.

	(1) SAR FE	(2) SAR RE
LR_Direct		
V-DEM Polyarchy	-2.17	-1.09
	(7.87)	(8.31)
Left-Wing Gov	15.6***	19.8***
	(1.81)	(1.89)
Change in GDP per Capita	-0.081	-0.13
	(0.19)	(0.20)
Trade as % of GDP	0.30***	0.28***
	(0.058)	(0.057)
International Migration Stock as % of Population	-0.72	0.063
	(1.01)	(0.92)
Refugees as % of Population	1.42	0.41
	(1.57)	(1.67)
Emigrants in US and Spain as % of Pop.	-0.70	0.47
	(0.59)	(0.53)
LR_Indirect		
V-DEM Polyarchy	0.54	-2.01
	(2.30)	(15.4)
Left-Wing Gov	-4.36***	35.6***
	(1.30)	(5.78)
Change in GDP per Capita	0.023	-0.24
	(0.057)	(0.37)
Trade as % of GDP	-0.084**	0.51***
	(0.028)	(0.13)
International Migration Stock as % of Population	0.20	0.12
	(0.30)	(1.67)
Refugees as % of Population	-0.39	0.76
	(0.46)	(3.01)
Emigrants in US and Spain as % of Pop.	0.21	0.79
	(0.19)	(0.91)
Observations	551	551
R^2	0.088	0.347

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Regulatory Complexity Spatial Panel Data Models. Total Effects.

	(1) SAR FE	(2) SAR RE
LR_Total		
V-DEM Polyarchy	-1.62	-3.10
	(5.67)	(23.7)
Left-Wing Gov	11.2***	55.5***
	(1.78)	(6.87)
Change in GDP per Capita	-0.058	-0.38
	(0.14)	(0.56)
Trade as % of GDP	0.21***	0.80***
	(0.046)	(0.18)
International Migration Stock as % of Population	-0.52	0.19
	(0.73)	(2.59)
Refugees as % of Population	1.03	1.17
	(1.15)	(4.68)
Emigrants in US and Spain as % of Pop.	-0.50	1.26
	(0.42)	(1.43)
Observations	551	551
R^2	0.088	0.347

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$