European Journal of Political Research **53**: 822–839, 2014
doi: 10.1111/1475-6765.12058

Contextual variation in interdependent policy making: The case of tax competition

FABIO WASSERFALLEN
Department of Political Science, University of Zurich, Switzerland

Abstract. Many studies show that policy makers react to the policy choices made in other jurisdictions, but we still know relatively little about the factors driving interdependent policy making, especially about how context shapes interdependence. Theoretical arguments suggest that contextual factors, such as stable institutions and geographic location, explain variation in interdependence. However, there is a lack of empirical research investigating contextual heterogeneity in interdependent policy making, mainly because it cannot be analysed with standard spatial econometric methods. This article introduces multilevel modeling that allows the study of contextual variation in interdependence and illustrates the method with the analysis of uneven tax competition in Switzerland. The findings of fine-grained data show that cantonal governments compete more strongly with their competitors the closer a unit is located to a metropolis with comprehensive public good provision. The analysis demonstrates that we can better understand the mechanisms of interdependent policy making by studying its contextual drivers.

Keywords: interdependent policy making; tax competition; contextual variation; multilevel modeling

Introduction

Research has shown that policy making is driven by interdependence – that is, policy makers react to the policy choices made in other jurisdictions. The literature has recently shifted towards the study of the mechanisms of interdependent policy making, building on seminal work that has defined broad classes of mechanisms, such as competition, learning and emulation (Simmons et al. 2008). These typologies are useful starting points, but they only group similar dynamics. The mechanisms of interdependence are more complex and need further theorising. Thus, despite advances in the literature, the main challenge, to which this study aims to contribute, remains to explore theoretically and empirically why policy makers react to the decisions made in other jurisdictions. The basic argument of this article is that we can better understand interdependent policy making by studying how contextual factors, such as stable institutions and geographic location, shape the extent to which policy makers react to the decisions of other governments.

Thus far, the conditionality of interdependent policy making has largely been ignored (Neumayer & Plümper 2012). The standard approach in the literature is that researchers specify spatial econometric models based on the assumption that interdependence is uniform among the units of analysis. Only few studies have gone beyond this homogeneity assumption by analysing variation in interdependent policy making (Basinger & Hallerberg 2004; Cao & Prakash 2012). What has been almost completely neglected is that there are contextual (i.e., time-persistent) factors, such as institutional and geographic variables, that explain variation in the extent to which policy makers react to the decisions made in other

jurisdictions.[1] For example, we might expect that countries with less veto institutions are more responsive to the decisions of other governments. That there is a lack of empirical research investigating the contextual conditionality of interdependence is mainly because it cannot be analysed with standard spatial econometric methods. I propose, as an alternative, multilevel modeling that allows us to study whether contextual variation between the units of analysis conditions the extent to which policy makers react to the policy choices made in other jurisdictions.

I illustrate the method and the study of contextual heterogeneity in interdependent policy making with a quintessential example of interdependence – namely tax competition. Building on the literature on uneven tax competition and spillover models (Bucovetsky 1991; Dehejia & Genschel 1998), I argue that sub-national units located at nearby metropolises attract a relatively large share of mobile taxpayers by offering low taxes and the prospect of living close to a dynamic region providing a comprehensive set of public goods. In addition, I hypothesise that the increased share of mobile taxpayers in these jurisdictions enhances the pressure on governments to keep taxes at competitive levels. Accordingly, the more closely a sub-national jurisdiction is located to a metropolis, the more strongly is its government expected to react to the tax policy changes of competitors. This hypothesis is analysed empirically with exceptionally fine-grained data for income taxation in Switzerland.

The empirical analysis shows that we can clearly reject the unrealistic assumption of homogeneous interdependence in Swiss tax competition. Swiss cantons react to varying extents to the tax policy changes of competitors. The findings support the argument that spillover from metropolises shapes tax competition dynamics. In the case of Swiss tax competition, variation in interdependence between the cantons is basically a function of their geographic locations. In other cases, however, the relevant heterogeneity explaining variation in responsiveness between the units of analysis are, for example, differences in economic factor endowments or domestic institutions (e.g., veto points). I discuss several studies that make the argument that some structural factors condition interdependent policy making to show that the study of contextual variation in interdependence is relevant beyond the case of Swiss tax competition. In sum, this analysis shows that we can elaborate a more nuanced understanding of the mechanisms of interdependent policy making by studying contextual variation in responsiveness with multilevel time-series cross-section analysis.

The article is structured as follows. It begins by situating the study in the literature on interdependent policy making and formulating the theoretical argument. It then presents the Swiss case and explores the empirical method before reporting the findings. Prior to presenting the conclusions, the relevance of the developed approach beyond the Swiss case is discussed.

Towards the study of contextual variation in responsiveness

In an ever more interconnected world, the argument that policy making is an interdependent phenomenon has become intensively studied in political science (Graham et al. 2013). The critical juncture triggering a vast and growing literature over recent years was the shift

from the rather narrow focus on the spread of (typically successful) policies towards the analysis of why policy makers are responsive to the decisions made in other units.

Seminal theoretical work has identified broad classes of mechanisms explaining why policy decisions are systematically conditioned by policy choices made in other jurisdictions – one of which is that governments compete for capital, taxpayers or investment (Simmons et al. 2008; Gilardi 2012). The basic idea is that tax, economic and regulatory policy making are – at least partly – a function of policy makers' reactions to the choices made in other jurisdictions that are considered to be competitors in attracting fiscal revenue and economic activity. The argument that policy makers are responsive to the decisions of competitors is for many policy areas a sound starting point for studying interdependent policy making. However, the competition mechanism is – like the other mechanisms of interdependence – only a general classification that groups dynamics following a similar general rationale. Despite useful definitions of mechanisms and the substantial research effort of the last years, we still know relatively little about the specific dynamics of interdependence. Thus, the main challenge of the literature, to which this study aims to contribute, remains to theoretically and empirically unpack interdependent policy making.

Spatial econometric modeling has become the standard empirical approach in the literature. Models using this method include, as their main explanatory variable, a so-called 'spatial lag', which is the weighted average of the dependent variable in other units. The estimates of the spatial lag reveal whether policy makers are responsive to policy choices made in other jurisdictions. The key connection between the theoretical turn to the study of mechanisms and the methodological advances in spatial analysis revolves around the specification of the connectivities. Scholars studying interdependence use various theory-driven connectivities, such as traffic data, network measures and economic indicators (Zhukov & Stewart 2013). Despite these advances, however, there is still some concern that spatial models are based on over-simplified assumptions (Ward & Grundig 2011).

One way researchers can elaborate more nuanced theoretical and empirical models of interdependence is by focusing on the conditionality of responsiveness. Only surprisingly few studies have done this. The vast majority of articles assume that interdependence is uniform between the units, although this is often an unrealistic assumption. Neumayer and Plümper (2012) list a series of theoretical arguments explaining variation in interdependence – for example, that democratic governments are more responsive to outside influences than autocratic ones.

So far, there are basically two empirical strategies in the literature for studying the conditionality of interdependent policy making. The first – a standard design in the yardstick competition literature – is to split the full dataset into two sub-samples, where the units of one sub-sample are expected to be more responsive than the units of the other sub-sample. Concretely, yardstick competition scholars test whether incumbents that are at risk of losing elections are more responsive to the tax decisions of their neighbours (compared to incumbents that cannot run for re-election or relatively safe government coalitions). The difference between the spatial lag estimates of the two sub-samples then reveals whether the empirics support the yardstick hypothesis (Besley & Case 1995; Elhorst & Fréret 2009). This is a reasonable strategy if the argument on variation in responsiveness distinguishes two well-specified groups.

The second approach is to investigate the conditionality of interdependence by interacting the spatial lag and a mediating variable using standard time-series cross-section models with unit fixed effects (Basinger & Hallerberg 2004; Cao & Prakash 2012). The consequence of controlling for unit heterogeneity with unit fixed effects is that the empirical analysis is limited to the study of the variance within the units (over time). In regard to the spatial lag, this is what researchers aim to model as they are typically interested in how policy changes in other jurisdictions affect policy making. Yet this is also of relevance – and may be problematic – for the conditioning effect. With this model specification, researchers are constrained to the analysis of how variation in responsiveness is mediated by temporal changes of a conditioning variable, which is fine if the theory predicts that temporal variation of economic, political or social factors within a society mediate interdependence. If, however, we hypothesise that contextual (i.e., rather time-invariant) differences between the units of analysis explain variation in responsiveness, we cannot rely on the unit fixed effects specification.

In other words, existing spatial econometric approaches either compare two contexts or investigate the mediating effect of a variable on responsiveness within countries (over time). As an alternative approach, I introduce multilevel modeling that allows researchers to study contextual variation in responsiveness. With multilevel models we can, for example, test how stable institutions and geographic locations condition the extent to which policy makers react to the policy decisions made by other units. As a concrete example, I explore whether the responsiveness of governments in tax competition depends on the geographic location of their jurisdictions. The advantage of multilevel modeling is not simply a technical issue, but one with substantive theoretical implications since the approach introduced in this article is targeted at investigating contextual variation in interdependence between the units of analysis (beyond the comparison of only two contexts).

Uneven tax competition

Taxation may be the most studied policy area within the recent upsurge of research on interdependent policy making, which is not surprising given the straightforward externalities of tax policy making in open societies (e.g., Basinger & Hallerberg 2004; Jensen & Lindstädt 2012).[2] Besides the research on international tax competition, there is a vast amount of theoretical work in the field of fiscal federalism. The most famous model in this literature is that of Tiebout (1956), according to which people move to jurisdictions that provide their preferred public goods most efficiently. Controversially discussed in the literature is whether tax competition enhances welfare, as the Tiebout (1956) model suggests, or whether tax competition leads to dysfunctional race-to-the-bottom dynamics (Oates 2001). Instead of analysing the overall aggregate effects of tax competition, the main motivation of this study is to explore variation in the extent to which governments engage in tax competition.

The mobility of the tax base is the basic factor triggering tax competition. Accordingly, jurisdictions compete with those jurisdictions to which their tax base might move when the tax rate differences change. This tells us with whom governments are potentially competing.

The key for explaining variation in the extent to which governments engage in tax competition, however, is to focus on what kind of heterogeneity among the jurisdictions affects the mobility of the tax base. Uneven tax competition models show that the governments of jurisdictions with specific contextual characteristics attract mobile taxpayers by setting low tax rates (Bucovetsky 1991). These jurisdictions accommodate a larger share of mobile taxpayers, which may make their governments more responsive to the tax policy changes of competitors because they want to ensure that their mobile tax base does not relocate (again). The textbook case illustrating uneven tax competition emphasises country size as a mediating contextual factor. Dehejia and Genschel (1998, 13) argue that small countries 'can exploit tax competition as a type of beggar-thy-neighbor policy' by attracting mobile taxpayers from large countries. The basic intuition of the argument is that small jurisdictions benefit from undercutting the tax rates of competitors because the revenue gains from the attracted (mobile) tax base outweigh the revenue losses from the already present tax base. Accordingly, small jurisdictions are expected to tax at lower levels than large ones.

Uneven tax competition models also predict that the structures of the tax bases change. More concretely, the share of mobile taxpayers is supposed to increase in small countries and to decrease in larger ones. As a consequence, 'the elasticity of the tax base with respect to the tax rate . . . is less strongly negative in large jurisdictions than in small jurisdictions' (Brülhart & Parchet 2014: 67–68). Stated in less technical terms, the taxpayers of small countries are expected to be more mobile – that is, they will relocate more quickly when the tax rates become more attractive in other units. If the tax bases of small jurisdictions are more sensitive to tax rate differentials, as suggested by uneven tax competition models, we would also expect that their governments are more sensitive to tax rate changes of competitors. Uneven tax competition models thus provide the theoretical grounds for the following hypothesis: governments of small countries are more responsive to the tax rate changes of competitors because they accommodate a larger relative share of mobile taxpayers.

Shifting from the international to the subnational level, the literature has identified another factor that may drive tax competition: the extent to which jurisdictions benefit from spillover of nearby jurisdictions. In a federal setting, public goods typically exert beneficial externalities beyond the sub-national units that provide them. Centres realise economies of scale and provide comprehensive public goods, such as cultural facilities, international airports, high-technology health care and major universities. All these goods exert positive externalities beyond their border, which gives a comparative advantage to nearby jurisdictions who benefit from the metropolises' public good provision without bearing the costs (Olson 1969; Hochman et al. 1995; Wilson 1999; Brueckner 2003).

Following the above discussion on the tax competition effects of country size, we can again formulate a hypothesis on variation in tax levels and one on tax responsiveness. First, the hypothesis on tax levels predicts that governments of the jurisdictions benefiting from the infrastructural supply provided by nearby metropolises tax at lower levels because they can offer low taxes and quick access to the economic and cultural activities of the metropolises. Second, the hypothesis on tax responsiveness suggests that governments of jurisdictions benefiting from spillover react more strongly to the tax rate changes of competitors because they attract tax-sensitive mobile taxpayers by offering low taxes and a comprehensive public good provision. The basic rationale of the main argument is that the

increased relative share of mobile taxpayers puts more pressure on governments to keep taxes at competitive levels. Thus, *governments react more strongly to the tax rate changes of competitors, the more a jurisdiction benefits from the positive externalities of nearby metropolises*. I test this hypothesis on variation in tax responsiveness for the case of income taxation in Switzerland, which provides an ideal setting to study sub-national tax competition.

Tax competition in Switzerland: System and data

Compared to other developed countries, Switzerland holds a top position in terms of the extent of fiscal decentralisation (OECD 2010). In particular, income and wealth taxation is predominantly sub-national. The federal government collects only a quarter of the overall tax revenue from direct taxation. The 26 cantons have different tax systems, and the cantonal taxation levels vary substantially, especially for high-income earners. For them, moving to a low-tax canton is fiscally attractive and feasible due to the comprehensive transportation system in the small country. Thus, a change of residence to a nearby canton with lower taxes is not necessarily associated with a change of workplace.

Relocation of famous wealthy people, such as top tennis player Roger Federer's move from Basel District to the low-tax canton of Schwyz, are usually accompanied by fierce debates on the pros and cons of tax competition. That cantons compete for wealthy taxpayers is an omnipresent political topic in Switzerland. A recent popular initiative from the Social Democrats, demanding a minimum sub-national tax rate for high-income earners, is an example of this debate.[3] That Swiss cantons do compete with one another is by and large taken for granted in Switzerland. The plausibility of the micro-foundation of tax competition makes the Swiss case an ideal setting to study the nuances of tax competition dynamics. In addition, the good overall comparability of the 26 cantons (compared to comparative cross-national studies) and the extraordinary data quality provide a solid basis for rigorous empirical analysis. In the following, I present the operationalisation of the dependent and the explanatory variables. For data availability reasons, the analysis is restricted to the period 1990–2009.

The dependent variable

I investigate inflation-adjusted effective income tax rates for a prototypical taxpayer.[4] Compared to other tax measures, such as average and statutory tax rates, this is an exceptionally fine-grained and accurate measure. I focus on high-income earners, who are the relevant tax base from a tax competition perspective.[5] The annual income analysed was roughly CHF150,000 (in 1990) and CHF200,000 (in 2009). (As of January 2014, €1 was equivalent to around CHF1.22.) As a reference, the median gross annual salary in 2008 was about CHF70,000. The effective tax rate is the sub-national tax due that these high-income earners pay as a share of their total incomes. Cantonal differences in tax systems and deductions are explicitly accounted for. In short, the tax rates used measure what a prototypical high-income taxpayer effectively transfers to the sub-national tax administration. The tax-level variation between cantons is large: the low-tax canton of Zug taxes the high-income earners under investigation at about 8 per cent, while they are taxed by more

than 18 per cent in the canton of Basel. Also, the variation within cantons over time is substantial (all cantons changed their tax rates several times from 1990 to 2009).[6]

The spatial lag variable

One of the main explanatory variables in this study is a spatial lag that accounts for competitive dynamics. The critical research step for scholars modeling interdependent policy making is the specification of the connectivity matrix ($\boldsymbol{W}$). In a row-standardised connectivity matrix, each cell of a row assigns to a specific canton the relative influence that the other cantons exert on it. Accordingly, a high value of w_{ij} means that canton j is an important competitor of canton i. In this analysis, the connectivity measure is supposed to account for the competitive tax pressure that the cantons exert on each other. What matters in our context is that cantonal policy makers assume that high-income earners move because of tax-level differences (which may be a wrong assumption, as the study of Brülhart and Parchet (2014) shows).

In the Swiss case it is uncontroversial to say that cantonal governments believe they are in competition with one another for high-income earners. But which cantons are considered to be competitors? Policy makers typically regard nearby jurisdictions, to which some of their taxpayers might easily move if tax differences change, as competitors (Feld & Reulier 2009; Gilardi & Wasserfallen forthcoming). Based on that proposition, I construct a connectivity matrix with exceptionally fine-grained travel data from the Federal Office for Spatial Development, assuming that the faster a canton can be reached, the higher the perceived competition pressure stemming from it. The data report the travel distance in minutes for each of the more than nine million dyads of 3,114 country-wide spatially distributed dots.[7] The connectivity measure is derived in three steps: first, I extract all the car travel distances that connect two cantons; second, I compute the mean travel time of all the extracted dyadic combinations of the travel distances between two cantons; and finally, I code in the connectivity matrix the row-standardised inverse of the estimated mean travel distance. In substantive terms, I assign higher weights the more closely two cantons are connected (measured in car travel time).[8]

The spillover variable

I hypothesised above that the geographic location of cantons explains variation in tax responsiveness. Cantons located close to metropolises have a comparative advantage in tax competition because they can benefit from public provisions, such as universities, hospitals and cultural facilities, without bearing the costs. The governments of these cantons attract a high share of mobile taxpayers by offering low tax rates and quick access to an economically and culturally vibrant metropolis. Because of the attracted higher share of mobile taxpayers, these cantons are more exposed to tax competition and their governments are thus expected to react more strongly to the tax rate changes of competitors. In short, the more a canton benefits from spillover, the more responsive will its government be to the tax rate changes of its competitors.

To measure the extent to which cantons benefit from the positive externalities of metropolises, I define which cantons qualify as metropolises and estimate the positive

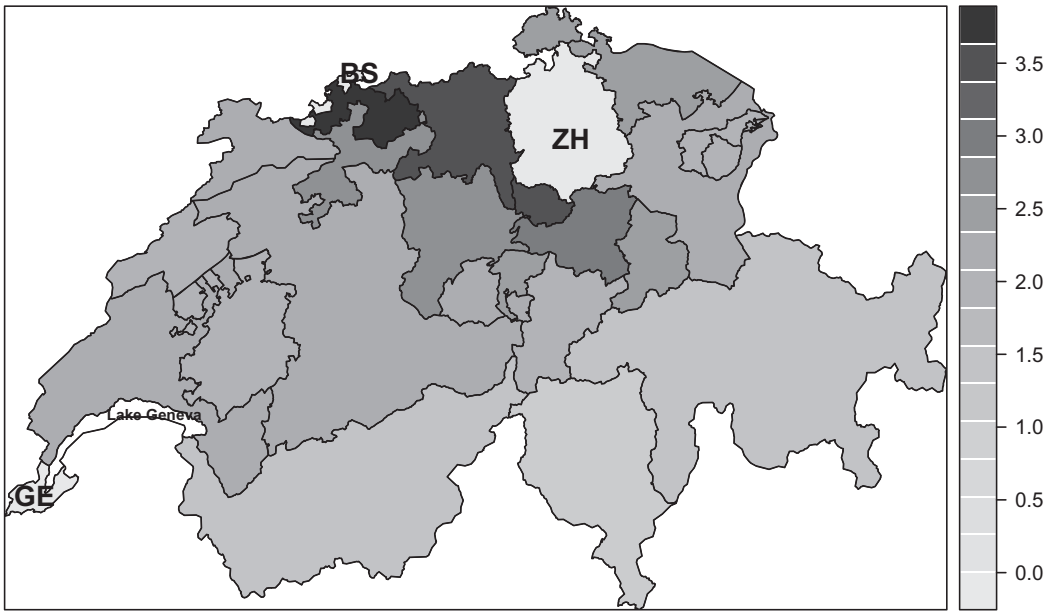

Figure 1. Map of the spillover variable.

externalities that they exert on the surrounding cantons. Based on a series of socioeconomic measures, the Swiss Federal Statistical Office categorises the canton of Zurich and the two city-cantons, Basel and Geneva, as the 'metropolitan areas' of Switzerland. These three economically vibrant cantons provide public goods with strong impacts beyond their borders. To measure the size of the positive externalities, I rely on the culture, leisure, education and health facility spending of these cantons; and, for estimating how much the surrounding cantons benefit from the positive externalities, I compute the travel time distances from the cantons to the metropolises. The spillover variable is then operationalised by estimating the product of the mentioned spending items (as a measure of the extent of the positive externalities) and the inverse of the travel distance to the metropolises (as a measure of the closeness to the metropolises). The substantive interpretation of the spillover variable is as follows: the faster the residents of a canton can reach metropolises with high spending in leisure, culture, education and health, the higher the spillover value is for that canton. Finally, I assign the value 0 to the three metropolitan cantons because they do not benefit from spillover of the other metropolises. Figure 1 maps the variation of the spillover variable between cantons. The cantons around Zurich (ZH) and Basel (BS) have the highest spillover values, while the rural cantons in the area of the Swiss Alps have low values.

Empirical model

To test whether the structural spillover variable mediates tax responsiveness as hypothesised, I rely on spatial multilevel modeling of time-series cross-section data. While

standard time-series cross-section models with unit fixed effects only analyse variation within units (over time), contextual differences between units can explicitly be modeled in a multilevel set-up. Accordingly, multilevel models allow researchers to explore how contextual variation in institutions or geographic location (as in our case) mediates the effects of time-series variables. Several scholars have outlined this powerful property of multilevel time-series cross-section models (Western 1998; Zorn 2001; Bafumi & Gelman 2006; Bartels 2008; Fairbrother 2014). Unfortunately, their advice has gone largely unnoticed. The key feature of multilevel time-series cross-section models is that the variations in the dependent variable are separated between units and over time.

A generic linear varying-intercept model of time-series cross-section data with a spatial lag is specified as follows:

$$\text{Level 1:} \quad y_{it} = \beta_{0i} + \beta_1 y^{w}_{it-1} + \rho \boldsymbol{w_i y^{w}_t} + \boldsymbol{\beta_2 x^{w}_{it}} + \varepsilon_{it} \tag{1}$$

$$\text{Level 2:} \quad \beta_{0i} = \gamma_0 + \gamma_1 z_i + \boldsymbol{\gamma_2 \bar{x}_i} + \mu_i \tag{2}$$

The dependent variable y_{it} – in this study the income tax rate in canton i at time t – is modeled as a linear function of a set of within-canton variables (on Level 1) and across-canton variables (on Level 2). Both equations include a stochastic term, ε_{it} and μ_i, which accounts for unobserved heterogeneity at both levels and follow a normal distribution ($\mu_i \sim N(0, \sigma_i^2)$; $\varepsilon_{it} \sim N(0, \sigma_{it}^2)$). The subscript i indicates that the intercept, β_{0i}, varies between cantons. This across-canton variation is modeled in the Level 2 equation.

The Level 1 equation, however, models tax variation over time. The superscript w indicates that the variables on Level 1 are transformed to within-canton predictors. The within-canton transformation is simply the deviation of a time-varying observation from the mean over the investigated time period (e.g., $x^{w}_{it} = x_{it} - \bar{x}_i$). The interpretation of the coefficients on Level 1 is like the interpretation of coefficients in standard time-series cross-section models with unit fixed effects: they report the average linear effect of a one-unit change in the explanatory variables on a one-unit change in the dependent variable. The within-canton transformed spatial lag is the main explanatory variable on Level 1 and accounts for competitive interdependence. The coefficient ρ is expected to be positive. Concretely, $\boldsymbol{w_i y^{w}_t}$ is the within-canton transformed spatial lag value for canton i at time t, whereas the vector $\boldsymbol{w_i}$ is row i of the connectivity matrix $\boldsymbol{W}$, and the vector $\boldsymbol{\gamma_t}$ is the dependent variable at time t. Following the literature, I model the spatial lag contemporaneously, assuming that governments interact strategically. Franzese and Hays (2007) show that simultaneity might bias the estimates, recommending instrumented spatial lags as a possible fix. The online appendix reports robust results for a model with an instrumented spatial lag.

To control for serial correlation, I include on Level 1 the within-canton transformed lagged dependent variable, y^{w}_{it-1}, (Keele & Kelly 2006).[9] As controls, I also include a vector of within-canton transformed variables, $\boldsymbol{x^{w}_{it}}$. I introduce the following fiscal, economic and political controls: the share of leftist seats in government, GDP per capita, population size and debt per capita. All else being equal, I expect that growing debts and populations, as

well as shifts to the left in the government, are associated with tax increases and that economic growth is negatively correlated with income tax levels.[10]

The Level 2 equation can be thought of as a simple cross-section linear model that explains the average taxation level of a canton over the investigated time period. The most important explanatory variable on Level 2 is the time-invariant spillover variable, z_i. I expect that cantons benefiting from spillover tax high-income earners at lower levels. To avoid biased estimates, I introduce the means of the controls, $\bar{\boldsymbol{x}}_i$ (Bafumi & Gelman 2006), whereas the average share of leftist seats in government is supposed to be positively correlated with tax levels, and smaller cantons are expected to tax at lower levels. I expect that lower levels of GDP per capita and higher debts are associated with higher income tax rates.

Note that the interpretation of the estimates of the variables is substantively different on both levels. For example, while the estimates of the average strength of the left in the government on Level 2 show whether cantons with an overall stronger leftist government participation tax high-income earners at higher levels, the estimates on Level 1 indicate whether shifts to the left in the government (over time) are associated with tax increases. By modeling the same variables on both levels we can assess their explanatory power as predictors of tax rate variation between cantons and over time.

The model discussed so far allows the intercepts to vary between cantons. The flexibility advantage of the multilevel approach is that we can straightforwardly extend the varying-intercept model to a model that allows the slopes (coefficients) to vary as well. Since we are interested in the heterogeneity of the cantons' reactions to the tax rate changes of competitors, ρ_i is estimated as a varying coefficient ($\rho_i = \gamma_0^{\rho} + \mu_i^{\rho}$, while $\mu_i^{\rho} \sim N(0, \sigma_i^2)$). The estimates of the varying-slope model show whether cantons are to varying extents responsive to competitors' decisions, providing evidence for the (im)plausibility of the assumption that all units react homogeneously to the decisions of other units.

Finally, I use the spillover variable, z_i, as a explanatory variable of ρ_i to test our main hypothesis that a government will react more strongly to the tax decisions of its competitors, the more a canton benefits from spillover. This cross-level interaction allows us to explore whether the spillover variable, as a contextual Level 2 variable, mediates tax responsiveness measured with the spatial lag (see Equation 5).

$$\text{Level 1:} \quad y_{it} = \beta_{0i} + \beta_1 y_{it-1}^{w} + \rho_i \boldsymbol{w_i y_t^w} + \boldsymbol{\beta_2 x_{it}^w} + \varepsilon_{it} \tag{3}$$

$$\text{Level 2:} \quad \beta_{0i} = \gamma_0 + \gamma_1 z_i + \boldsymbol{\gamma_2 \bar{x}_i} + \mu_i \tag{4}$$

$$\text{Varying spatial lag coefficients:} \quad \rho_i = \gamma_0^{\rho} + \gamma_1^{\rho} z_i + \mu_i^{\rho} \tag{5}$$

In the following empirical analysis, I report the estimates of three models. Model 1 is the varying-intercept model (see Equations 1 and 2). Model 2, the varying-slope model, includes a random effect for the spatial lag to explore whether there is variation in interdependence between cantons. Finally, the cross-level interaction model, model 3, extends model 2 by introducing the spillover variable as a predictor of the varying spatial lag coefficients (see Equations 3, 4 and 5). The estimates of model 3 indicate whether the

spillover variable is a good explanatory variable of the variation in tax responsiveness. All models are estimated with maximum likelihood estimation using the arm package in R (Gelman & Su 2013).

Empirical findings

Table 1 reports the empirical findings. Let us begin with the Level 2 variables that explain tax variation between cantons. The estimates of the spillover variable show that the more a canton benefits from the positive externalities of nearby metropolises, the lower the tax levels are for high-income earners. Also, as expected, cantons that had – over the investigated time period – stronger government participation by the left, tax high-income earners at higher levels. Furthermore, the findings on Level 2 indicate that cantons with a higher GDP per capita have leverage to tax at lower levels, while higher debt burdens are associated with higher taxes.

On Level 1, where tax rate variation over time is modeled, the most important explanatory variable is the spatial lag. The estimates of the spatial lag, which account for the responsiveness of policy makers to the tax rate changes of competitors, document a very substantial effect that is estimated with great precision. This supports the argument that cantonal tax decision making is strongly driven by tax competition. Changes in the debt levels, the government composition and the population size are, however, not systematically associated with tax rate changes. GDP per capita is the only control variable on Level 1 with a statistically significant negative correlation.

Comparing the estimates on Levels 1 and 2 illustrates the nuanced empirical testing that multilevel modeling allows. For example, the findings of the variables on the government participation of the left on Level 1 suggest that a shift within a canton to a more leftist government is not systematically associated with tax rate increases, while the Level 2 estimates indicate that cantons that had – over the investigated time period – stronger government participation by the left, tax at higher rates. This makes sense in the Swiss case. Shifts in government composition do not lead to dramatic political changes because cantonal governments are typically consensual multiparty ones held accountable by direct democratic institutions.

Finally, and most importantly, the findings show that there is substantial variation in responsiveness between cantons. Model 1 assumes homogeneous responsiveness, while model 2 adds a random effect for the spatial lag to allow for varying spatial lag coefficients. The large drop in the deviance (a standard statistical summary of model fit) from 931.8 in model 1 to 915.1 in model 2 documents the sharp improvement in model fit after allowing for varying responsiveness. Based on that analysis, we can clearly reject the assumption of homogeneous responsiveness. Furthermore, model 3 extends model 2 by introducing a cross-level interaction, which not only lets the spatial lag coefficients vary between cantons, it additionally models the spillover variable as an explanatory variable of the variation in tax responsiveness. The findings indicate that the extent to which cantons benefit from spillover indeed mediates tax responsiveness: the higher the spillover, the stronger the reaction of policy makers to the tax rate changes of competitors. Given that model 2 is nested in model 3, we can again compare the model fit to analyse the statistical significance of the cross-level

Table 1. Hierarchical time-series cross-section models explaining sub-national income tax rates for high-income earners: varying-intercept (model 1), varying-slope (model 2), and cross-level interaction (model 3)

	(1)	(2)	(3)
Level 2 (variation between units)			
Intercept	18.540*** (2.064)	18.554*** (1.824)	17.868*** (1.852)
Spillover (z)	–0.965* (0.543)	–1.241*** (0.479)	–0.907* (0.509)
GDP per capita ($\bar{x}_1$)	–1.023*** (0.299)	–0.953*** (0.263)	–0.959*** (0.262)
Leftist government participation ($\bar{x}_2$)	0.155*** (0.045)	0.169*** (0.040)	0.169*** (0.039)
Debt per capita ($\bar{x}_3$)	0.187* (0.109)	0.184* (0.097)	0.189* (0.097)
Population size ($\bar{x}_4$)	–0.128 (0.154)	–0.170 (0.135)	–0.170 (0.135)
Level 1 (variation over time)			
Lagged dependent variable (y^w_{t-1})	0.687*** (0.028)	0.662*** (0.027)	0.663*** (0.027)
Spatial lag (Wy^w_t)	0.484*** (0.050)	0.506*** (0.070)	0.246 (0.152)
GDP per capita (x^w_{1t})	–0.154*** (0.051)	–0.149*** (0.050)	–0.149*** (0.050)
Leftist government participation (x^w_{2t})	0.080 (0.329)	0.066 (0.326)	0.039 (0.326)
Debt per capita (x^w_{3t})	–0.003 (0.013)	–0.000 (0.014)	0.004 (0.014)
Population size (x^w_{4t})	0.010 (0.150)	–0.033 (0.146)	–0.047 (0.146)
Cross-level interaction			
Spillover * Spatial lag			0.123* (0.065)
Deviance	931.8	915.1	911.4
N (20 years * 26 cantons)	520	520	520
Varying intercepts	Yes	Yes	Yes
Varying spatial lag coefficients	No	Yes	Yes
Cross-level interaction	No	No	Yes

Note: Reported are the point estimates, the standard errors in parentheses and the significance levels (*** p > 0.01; ** p > 0.05; * p > 0.1).

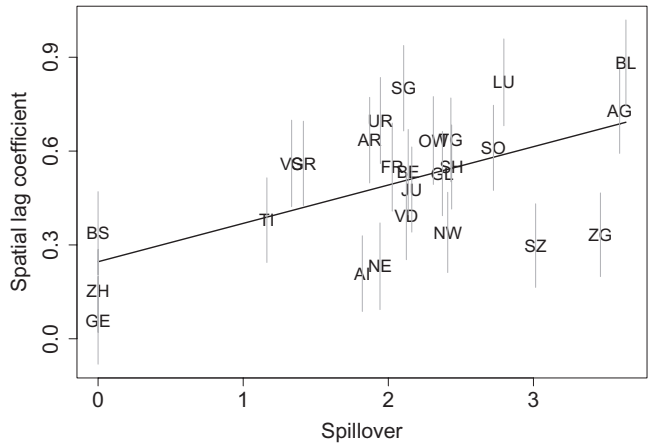

Figure 2. Varying spatial lag coefficients ±1 standard deviation plotted versus the spillover variable.

Notes: All estimates are based on model 3 of Table 1. Abbreviations (listed in the order of the federal constitution): ZH (Zurich), BE (Bern), LU (Lucerne), UR (Uri), SZ (Schwyz), OW (Obwalden), NW (Nidwalden), GL (Glarus), ZG (Zug), FR (Fribourg), SO (Solothurn), BS (Basel), BL (Basel District), SH (Schaffhausen), AR (Appenzell Outer-Rhodes), AI (Appenzell Inner-Rhodes), SG (St Gallen), GR (Grisons), AG (Aargau), TG (Thurgau), TI (Ticino), VD (Vaud), VS (Valais), NE (Neuchatel), GE (Geneva) and JU (Jura).

interaction. The deviance drops from 915.1 in model 2 to 911.4 in model 3, and the likelihood ratio test, comparing both models, estimates a p-value of 0.054, which shows that the cross-level interaction adds substantial explanatory power to the model. Overall, the estimates of the model support the argument that spillover explains variation in tax competition.

However, the tax responsiveness of some outlier cantons deviates from what the spillover argument predicts (inflating the standard error of the cross-level interaction). Figure 2 plots the correlation between the spillover variable and the varying spatial lag coefficients so as to visualise the cross-level interaction effect. The figure shows that the spillover variable is, overall, a good predictor of the variation in responsiveness, yet the spatial lag coefficients of a few cantons are not well explained by the spillover argument. The cantons of Schwyz (SZ) and Zug (ZG), for example, respond less intensively to the tax rate changes of competitors than expected. Policy makers in Zug and Schwyz may be confident that they can keep mobile taxpayers in their cantons because of the very low tax levels. The small rural canton of Appenzell Inner-Rhodes (AI) is another outlier with lower-than-expected tax responsiveness, which might be due to its distinct direct democratic political institutions.

In sum, the findings show that there is large variation in tax responsiveness among Swiss cantons and that part of that heterogeneity is because cantons compete on an uneven playing field.

Relevance beyond the Swiss case

I have introduced multilevel modeling to study whether contextual factors condition the extent to which governments react to the policy decisions made in other jurisdictions, and

I have illustrated the method with the case of uneven tax competition in Switzerland. I will now discuss several examples covering other policy areas and mechanisms of interdependence to show that the study of contextual variation in interdependent policy making is of relevance beyond the Swiss case. A rather straightforward extension of the present study to the international level builds on the argument that small countries compete more intensively for mobile tax sources than large ones (Bucovetsky 1991; Dehejia & Genschel 1998). So far, this proposition has been supported empirically by showing that small countries tax corporations at lower tax levels (Ganghof & Genschel 2008), which is, however, not a direct test of tax competition dynamics. The multilevel set-up used in this study allows the testing of whether the variation in population size between countries explains heterogeneity in the extent to which governments react to the tax changes of competitors.

Apart from geographic and demographic factors, domestic political institutions are another important variable conditioning competition among governments. Cao and Prakash (2012), for example, argue that veto players mediate the extent to which countries react to the decisions of trade competitors in environmental regulatory competition. We might similarly expect that the responsiveness of governments is higher in countries with majoritarian systems compared to consensus democracies. Using the unit fixed effects specification, Cao and Prakash (2012) investigated whether changes over time (within countries) in the veto player variable condition environmental regulatory competition. Their findings show that decreasing political constraints are associated with increasing responsiveness of governments to environmental regulatory changes of trade competitors. With the introduced multilevel approach we can analyse how institutional variation between countries mediates responsiveness (rather than how institutional variation over time conditions interdependence), which is more appropriate if we want to explore the variation of sticky political institutions between countries.

Finally, the study of contextual conditionality is not restricted to the competition mechanism and geographical or institutional arguments. For example, Lloyd et al. (2012) show that issue framing is crucial in the proliferation of laws against human trafficking. They argue that governments are more likely to criminalise human trafficking when it is linked to transnational crimes, such as drug trafficking, money laundering and smuggling. Building on their theory, we could expect that not all countries are equally exposed to the potential negative externalities of transnational crimes. An interesting extension of their study could thus investigate the variation between governments in their responsiveness to the criminalisation of human trafficking. Following the analysis of Lloyd et al. (2012), a government is more likely to adopt policies against human trafficking the more it is confronted with the problems of transnational crimes.

In sum, this discussion of different examples shows that the analysis of contextual variation in responsiveness has the potential to provide us with a more nuanced understanding of interdependent policy making for various mechanisms and policy areas.

Conclusion

The empirical and theoretical unpacking of why policy makers react to the decisions made in other jurisdictions is a challenging research task. Despite the vast amount of recently

published work on that topic, we still know relatively little about the specific dynamics driving interdependent policy making. I argue that focusing on contextual variation in responsiveness between the units of analysis is a promising strategy for shedding light on the black box of interdependence. The analysis of how, for example, institutional and geographic factors explain the extent to which some governments react more strongly to the decisions of others brings about a more nuanced picture of the mechanisms of interdependent policy making. In this article, I have discussed several examples covering different policy areas and mechanisms of interdependence to document the broader relevance of the study of contextual variation in responsiveness.

I also introduced multilevel modeling that allows the empirical analysis of the mediating effect of contextual (i.e., time-persistent) factors, such as institutional and geographic variables, on interdependent policy making. I have illustrated the method by investigating a typical case of how contextual variation between units shapes interdependence – namely uneven tax competition for high-income earners in Switzerland. Building on uneven tax competition and spillover models, I argued that cantons located near to metropolises attract a relatively large share of mobile taxpayers by offering low taxes and the prospect of living close to metropolises that provide a comprehensive set of public goods. As a consequence of the increased share of mobile taxpayers, these cantons are more exposed to tax competition and their governments can thus be expected to be more responsive to the tax policy changes of competitors. The analysis of exceptionally fine-grained data supports the hypothesis that variation in the extent to which cantons benefit from the spillover of nearby metropolises explains variation in tax responsiveness.

I believe that the study of the contextual conditionality of interdependence has the potential to advance the literature by elaborating a more nuanced understanding of the mechanisms of interdependent policy making. Apart from its importance for the academic literature, the study of contextual variation in responsiveness is also of practical importance. The better political actors understand which jurisdictions react more (or less) strongly to the decisions of other jurisdictions, the more effectively can they promote the spread of specific reforms or – depending on their political interests – constrain the dynamics of interdependent policy making. In the case of Swiss tax competition the findings of this study suggest that the containment of spillover effects, which could be achieved by shared financing of public goods, would mitigate tax competition dynamics.

Acknowledgements

I would like to thank Xun Cao, Johan E. Elkink, Katharina Füglister, Fabrizio Gilardi, Silja Häusermann, Lucas Leemann, Beth Simmons, Gabi Spilker, Marco Steenbergen, Julian Wucherpfennig, the *EJPR* editors and the anonymous reviewers for helpful comments, as well as Gema Ricart (Swiss Federal Tax Administration), Martin Tschopp (Federal Office for Spatial Development), Urban Roth (BAK Basel Economics) and Lukas Schmid (University of St Gallen) for sharing data. All remaining errors are my own. The bulk of this research was carried out while I was a visitor at the Weatherhead Center for International Affairs at Harvard University. I would like to express my gratitude to Beth Simmons, Steve

Bloomfield and Michelle Eureka for their support. The financial help of the Swiss National Science Foundation (Grant PBZHP1-138702) is also acknowledged.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Online appendix: Contextual variation in interdependent policy making: The case of tax competition
Table 1: Robustness checks
Table 2: Summary statistics

Notes

1. An exception is the diffusion literature using event history analysis. Arguments on the context conditionality are among the most interesting contributions of that literature (e.g., Shipan & Volden 2006).
2. For a helpful overview, see Genschel and Schwarz (2011).
3. On 28 November 2010, the electorate rejected this proposal with an overall disapproval rate of 58.5 per cent.
4. Gilardi and Wasserfallen (forthcoming) investigate 15 income categories that are, however, not adjusted for inflation. Inflation-adjusted tax data are not available for so many income categories.
5. The data are calculated for a married person without children. Tax deductions for children can make a difference, but that carries weight only for middle- and low-income families. The data are kindly provided by the Swiss Federal Tax Administration.
6. See the online appendix for detailed descriptive statistics.
7. Data are available for travel by car and public transport in 2005. Both datasets are highly correlated. I use car travel data because reports from the Federal Office for Spatial Development show that Swiss people use a car for more than two-thirds of the covered travel distances. Many thanks to the Federal Office for Spatial Development for sharing data.
8. Choosing other estimates than the mean only marginally changes the empirical findings (see the online appendix for robustness checks).
9. Introducing the lagged dependent variable solves the problem of serial correlation. While the residuals are strongly serially correlated in the model without the lagged dependent variable (p-value = 0.000), they are not systematically correlated after including the lagged dependent variable (p-value = 0.477).
10. Results are robust when the control variables are lagged by one year.

References

Bafumi, J. & Gelman, A. (2006). Fitting Multilevel Models When Predictors and Group Effects Correlate. Paper presented at the annual meeting of the Midwest Political Science Association, Chicago, IL, 20–23 April.

Bartels, B.L. (2008). Beyond Fixed versus Random Effects: A Framework for Improving Substantive and Statistical Analysis of Panel, Time-series Cross-sectional and Multilevel Data. Mimeo, George Washington University.

Basinger, S.J. & Hallerberg, M. (2004). Remodeling the competition for capital: How domestic politics erases the race to the bottom. *American Political Science Review* 98(2): 261–276.

Besley, T. & Case, A. (1995). Incumbent behavior: Vote-seeking, tax-setting and yardstick competition. *American Economic Review* 85(1): 25–45.

Brueckner, J.K. (2003). Strategic interaction among governments: An overview of empirical studies. *International Regional Science Review* 26(2): 175–188.

Brülhart, M. & Parchet, R. (2014). Alleged tax competition: The mysterious death of inheritance taxes in Switzerland. *Journal of Public Economics* 111: 63–78.

Bucovetsky, S. (1991). Asymmetric tax competition. *Journal of Urban Economics* 30: 167–181.

Cao, X. & Prakash, A. (2012). Trade competition and environmental regulations: Domestic political constraints and issue visibility. *Journal of Politics* 74(1): 66–82.

Dehejia, V.H. & Genschel, P. (1998). Tax competition in the European Union. *MPIfG Discussion Papers* 98(3): 1–40.

Elhorst, J.P. & Fréret, S. (2009). Evidence of political yardstick competition in France using a two-regime spatial Durbin model with fixed effects. *Journal of Regional Science* 49(5): 931–951.

Fairbrother, M. (2014). Two multilevel modeling techniques for analyzing comparative longitudinal survey datasets. *Political Science Research and Methods* 2(1): 119–140.

Feld, L.P. & Reulier, E. (2009). Strategic tax competition in Switzerland: Evidence from a panel of the Swiss cantons. *German Economic Review* 10(1): 91–114.

Franzese, R.J. & Hays, J.C. (2007). Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data. *Political Analysis* 15: 140–164.

Ganghof, S. & Genschel, P. (2008). Taxation and democracy in the EU. *Journal of European Public Policy* 15(1): 58–77

Gelman, A. & Su, Y.-S. (2013). arm: Data Analysis Using Regression and Multilevel Models, Version 1.6-09. Available online at: www.cran.r-project.org/web/packages/

Genschel, P. & Schwarz, P. (2011). Tax competition: A literature review. *Socio-Economic Review* 9(2): 339–370.

Gilardi, F. (2012). Transnational diffusion: Norms, ideas and policies. In W. Carlsnaes, T. Risse & B.A. Simmons (eds), *Handbook of international relations*, 2nd edn. London: Sage.

Gilardi, F. & Wasserfallen, F. (forthcoming). How socialization attenuates tax competition. *British Journal of Political Science*.

Graham, E.R., Shipan, C.R. & Volden, C. (2013). The diffusion of policy diffusion research in political science. *British Journal of Political Science* 43(3): 673–701.

Hochman, O., Pines, D. & Thisse, J.-F. (1995). On the optimal structure of local governments. *American Economic Review* 85(5): 1224–1240.

Jensen, N.M. & Lindstädt, R. (2012). Leaning right and learning from the left: Diffusion of corporate tax policy across borders. *Comparative Political Studies* 45(3): 283–311.

Keele, L. & Kelly, N.J. (2006). Dynamic models for dynamic theories: The ins and outs of lagged dependent variables. *Political Analysis* 14(2): 186–205.

Lloyd, P., Simmons, B.A. & Stewart, B.M. (2012). The Global Diffusion of Law: Transnational Crime and the Case of Human Trafficking. Working paper. Indiana University and Harvard University.

Neumayer, E. & Plümper, T. (2012). Conditional spatial policy dependence: Theory and model specification. *Comparative Political Studies* 45(7): 819–849.

Oates, W.E. (2001). Fiscal competition or harmonization? Some reflections. *National Tax Journal* 54(3): 507–512.

Olson, M. Jr. (1969). The principle of 'fiscal equivalence': The division of responsibilities among different levels of government. *American Economic Review* 59(2): 479–489.

Organisation for Economic Cooperation and Development (OECD) (2010). *Revenue Statistics 2010*. Paris: OECD.

Shipan, C.R. & Volden, C. (2006). Bottom-up federalism: The diffusion of antismoking policies from US cities to states. *American Journal of Political Science* 50(4): 825–843.

Simmons, B.A., Dobbin, F. & Garrett, G. (eds) (2008). *The global diffusion of markets and democracy*. Cambridge: Cambridge University Press.

Tiebout, C.M. (1956). A pure theory of local expenditures. *Journal of Political Economy* 64(5): 416–424.
Ward, H. & Grundig, F. (2011). Networks, Policy Diffusion and Chaotic Dynamics: Towards an Agent-based Model. Paper presented at the general conference of the European Political Science Association, Dublin, 16–19.
Western, B. (1998). Causal heterogeneity in comparative research: A Bayesian hierarchical modelling approach. *American Journal of Political Science* 42(4): 1233–1259.
Wilson, J.D. (1999). Theories of tax competition. *National Tax Journal* 52(2): 269–304.
Zhukov, Y.M. & Stewart, B.M. (2013). Choosing your neighbors: Networks of diffusion in international relations. *International Studies Quarterly* 57(2): 271–287.
Zorn, C. (2001). Estimating between- and within-cluster covariate effects, with an application to models of international disputes. *International Interactions* 27: 433–455.

Address for correspondence: Fabio Wasserfallen, Department of Political Science, University of Zurich, Affolternstrasse 56, 8050 Zurich, Switzerland. E-mail: wasserfallen@ipz.uzh.ch; URL: www.fabiowasserfallen.ch