Oxford University
Centre for
Business Taxation

Learning and international policy diffusion: the case of corporate tax policy

November 2013

WP 13/19

Johannes Becker
University of Münster

Ron B Davies
University of Dublin

Working paper series | 2013

SAID BUSINESS SCHOOL
UNIVERSITY OF OXFORD

Learning and International Policy Diffusion – The Case of Corporate Tax Policy

Johannes Becker
University of Münster
Institute for Public Economics
Wilmergasse 6-8
48143 Münster
Germany
johannes.becker@wiwi.uni-muenster.de

Ronald B. Davies
University College Dublin
G215 Newman Building
Belfield
Dublin 4
Ireland
ronbdavies@gmail.com

June 13, 2013

Abstract

A recent empirical literature has arisen documenting the response of one nation's policy choices, including tax, environmental, and labour policies, to that of others. This has been largely interpreted as evidence of competition, be it for mobile resources (like FDI, taxable book income etc.) or yardstick. We present a third explanation based on learning. When countries' tax choices reflect private information about unobserved conditions, this encourages nations to update their policies not in order to retain investment or manipulate trade flows, but because the new information conveyed by overseas tax rates allows them to fine-tune their own policies. With this "social learning", countries converge on their optimal policies faster than in isolation. Furthermore, this convergence implies a pattern of policy convergence often attributed to competition for mobile resources. The speed of this convergence is smaller in the presence of policy adjustment costs although it remains faster than convergence in isolation. In addition, adjustment costs result in inefficient policy adjustment because countries do not internalize the benefits conveyed by their own adjustments to other nations. Finally, we show that these baseline results are robust to alternative network architectures, the choice of which can be used to replicate stylized facts found in the empirical tax competition literature.

JEL classification: H25, H32, H87

Keywords: social learning, tax competition

1 Introduction

Recently, a body of empirical literature has arisen documenting an interdependence among countries' tax policies (see, e.g., Devereux et al. 2008, Overesch and Rincke 2009, 2011, Heinemann et al. 2010), labour policies (e.g. Davies and Valammenati, 2013, Olney 2010) and environmental policies (e.g. Davies and Naughton, forthcoming, Beron, Murdoch, and Vijverberg, 2003, and Murdoch, Sandler, and Vijverberg, 2003). This work generally finds that the policies in one country are positively correlated with the policies set elsewhere. The primary conclusion is that this is evidence of competition in the tradition of Zodrow & Mieszkowski (1986) and Wilson (1986) wherein governments compete for mobile resources by implementing "business-friendly" policies.[2] An alternative explanation is the yardstick competition of Besley and Case (1995) in which voters judge the performance of local policy makers by the policies implemented elsewhere.[3] In both of these, the rationale behind international policy diffusion is that the policy choices undertaken in one location affect the payoffs from a given policy in other locations.

In this paper, we formalize a third alternative in which the preferred policy choice in one country is affected by those elsewhere without relying on changes in payoffs, i.e. without hinging on cross-border capital flows (as in tax competition) or political competition under asymmetric information (as in yardstick competition). We do so by assuming that countries only have incomplete information on the true state of nature but can learn about it by observing the policies set elsewhere. This learning then results in policy convergence, not because the taxes chosen elsewhere affect the payoffs of a given nation's policy, but because learning results in the reduction in information asymmetries. Our model thus picks up the perspective taken by Slemrod (2004), where he documents both the decline in both corporate tax rates and the international variation of those rates.

We start the analysis by reviewing the empirical literature on policy-making in an international setting.[4] We do this in order to develop a set of overarching patterns in the data, in particular with regards to the correlation in policies across borders. In addition, this review of the common methodologies and estimation strategies gives us a framework through which we can compare the results from our model to these empirical regularities.

In a second step, we lay out our theory. We build a model where countries set tax rates that optimally depend on a common state of nature which is *a priori* unknown. In each period, each country receives a payoff that is a function of their tax policy, the state of nature, and a random variable. This payoff then

[2] It should be noted, however, that there also exist a smaller group of theories in which, due to negative externalities, governments compete to drive FDI to other nations. The study by Markusen, Morey and Oleweiler (1995) is an example of this "not-in-my-backyard" literature.

[3] See Salmon (1987) for an initial application to taxes and Brueckner (2003) for an overview.

[4] Although our discussion focuses on international policy diffusion, our theory equally applies to the literature on policy competition between jurisdictions within a nation. Empirical work here includes Fredricksson, List, and Millimet, 2003, Levinsohn, 2003, Heyndels and Vuchelen, 1998, Besley and Case, 1995, and Mintz and Smart, 2004.

acts as a private signal revealing information about the state of nature. In addition, in each period each country observes the taxes set elsewhere in the previous period which contain information about the (lagged) beliefs on the state of nature held elsewhere. As this information diffusion results in belief convergence, so too will it lead to tax convergence. More concretely, countries will choose to reduce their tax rate in the second round if the average first round tax rate is below its own tax rate and vice versa. The resulting pattern then appears as if tax rates were strategic complements and could, thus, be misinterpreted as evidence for tax competition.

Our approach builds on the social learning literature (Gale, 1996, Gale & Kariv, 2003, Vives, 1996) which assumes that agents can observe their neighbours' actions but not their outcomes.[5] We translate this theory into the realm of policy-making jurisdictions and reinterpret the evidence accordingly. Moreover, we contribute to this literature by accounting for policy-specific aspects (such as ideology) that affect the informational content of another country's actions.

In addition, we introduce adjustment costs, that is, a cost to changing the tax rate from the one that was used in the previous period. An important implication of positive adjustment costs is that it results in an inefficient equilibrium, an inefficiency that extends in general to any setting of social learning with adjustment costs. This inefficiency is very different from the fiscal externality that arises in the tax competition literature. There, because one nation does not internalize the impact of its tax rate on the tax base elsewhere (due to, for example, capital mobility), this typically results in equilibrium taxes which are inefficient. In particular, the equilibrium is often characterized by inefficiently low taxes and the underprovision of public goods. Here, however, the externality arises because although non-adjustment provides some information to other countries about a given nation's beliefs, it provides less information than actual adjustment. As the additional information from adjustment can help to refine the beliefs and therefore choices of other nations, there is an uninternalized positive information externality. A consequence of this is that adjustments happen inefficiently rarely. Thus, just as our model can provide a rationale for correlated tax rates across borders even without international linkages in real activity, it can also result in inefficiencies in international tax setting.

Note, however, that we are not attempting to suggest that tax competition for mobile capital or yardstick competition do not exist. Rather, we believe that the mechanisms in our model complement those results and provide a richer framework for the discussion of international policy diffusion. In particular, as tax convergence in our model is the result of beneficial information diffusion, our model provides a rationale for why the observed patterns in tax rates may be less damaging than what is typically supposed.[6]

The remainder of the paper is organized as follows. Section 2 briefly surveys

[5]In contrast, the social *experimentation* literature (see, for instance, Bala & Goyal, 1998) assumes that agents can observe their neighbours' actions *and* their outcomes.

[6]Other models, including Davies, 2005, provide additional settings where inter-jurisdictional tax competition can be welfare improving.

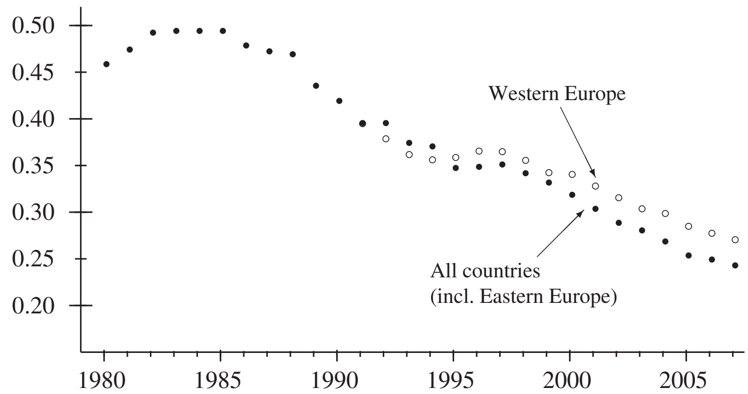

Figure 1: Statutory tax rates in European countries, 1980-2007 (Source: Heinemann et al. 2010)

the main findings of the empirical literature and discusses them from the theoretical point of view. Section 3 presents the model, derives the main results and discusses some extensions. Section 4 concludes.

2 Literature: Theory and evidence on corporate tax policy in open economies

In this section, we review the empirical literature on international tax competition (and international policy competition in general). Our goal here is two-fold. First, we seek to develop a set of empirical regularities into which a model of policy diffusion should be able to provide insight. Second, by reviewing the methodologies used in the empirical literature, it will ease our discussion of how our model can provide comparable patterns to what is observed in the data.

Regularity #1: Persistent secular downward trend in corporate tax rates. As discussed in detail in Slemrod (2004), starting in the 1980s, there has been a persistent downward trend in corporate tax rates in nearly all OECD countries as illustrated in Figure 1. The literature provides us with four alternative explanations for interpreting this trend.

First, the most popular approach is to point at the abolishment of capital controls in the 1980s which are supposed to have opened the gate to tax competition. Presuming that mobile capital is driven off by relatively high tax rates (something confirmed by the bulk of the literature, with de Mooij and Ederveen (2008) providing an overview), the resulting competitive pressure presumably led to the downward trend in corporate tax rates.

Second, the reason for the common trend in tax rates could be yardstick competition. As formalized by Besley and Case (1995), this is a situation in which the political success of a politician depends on the attractiveness of their policies relative to the policies in place elsewhere. When some countries demonstrate that providing public goods is possible with lower corporate tax rates, this is observed by other countries' electorates who then pressure their government to follow suit. This then could also result in the observed tax trends.

Third, and building off of either of the first two explanations, is that there has been a policy shift by a "leader" country such as the US. As is well-known, the US reduced its taxes between 1986 and 1988 and other countries followed suit. As first expressed by Gordon (1992), this interaction is strengthened by the US's role as the largest and most important capital exporter. Under the tax credit system, countries hosting US FDI with tax rates between the pre-reform and post-reform US rates would for the first time find that US investment was responsive to their tax. Accordingly, a drop in the US corporate tax rate could make a reduction in host countries' tax rates necessary.[7] In addition, there is the possibility that during the mid-1980s there was a change which gave the US leadership status in the Stackelberg sense, that is, led it to use its own tax policy to strategically manipulate those elsewhere. This change may have then resulted in a shift in US taxes as it sought to exploit this new advantage, something which then filtered through the follower countries.

Finally, the common downward trend could be due to common movement in the national characteristics that drive tax setting. For example, it could be the result of a common intellectual or ideological shift across OECD countries in favour of a lower tax rate regime. Indeed, such a possibility seems in line with the fact that the political landscape during the 1980s was dominated by leaders such as Margaret Thatcher, Ronald Reagan, and Helmut Kohl; all conservatives who may have shaped their societies and policies towards a more business-oriented attitude.

Thus, the existing literature by and large attributes the change in tax rates to either some form of competition (where the payoffs to a tax rate in one country depends on those elsewhere) or coincidence driven by a common trend in determining factors.

Regularity #2a: Countries react to their neighbours' tax rate changes. The most recent (and most sophisticated) empirical studies find that an individual country's tax policy reacts to tax changes in other countries, see Devereux et al. (2008), Davies & Voget (2009), Overesch & Rincke (2011) etc. The results from this literature suggests that countries can be observed to cut their tax rates *because* their neighbours did so.

For later use, it is worthwhile to have a closer look at the primary methodology in this literature which has been a spatial econometric approach.[8] In these

[7] Whalley (1990) considers the first two alternative approaches and concludes that neither of them is able to account for the similarities in the tax changes of the countries under consideration.

[8] For a complete treatment, we defer the interested reader to Anselin (1988).

works, the general regression specification is:

$$T_{i,t} = X_{i,t}\beta + \rho \sum_{j \neq i} \omega_{j,i,t} T_{j,t} + \varepsilon_{i,t} \quad (1)$$

where $T_{i,t}$ is a policy by country i (be it tax, environmental, labor, or other policy), $X_{i,t}$ is a vector of characteristics specific to country i (which can include lagged values and/or information on other countries), $\epsilon_{i,t}$ is an error term, and $\sum_{j \neq i} \omega_{j,i,t} T_{j,t}$ is known as the spatial lag. The spatial lag is the weighted average of the policies across other countries.[9] It aggregates the policies of other countries into a single variable using a weighting scheme where $\omega_{j,i,t}$ is the weight that i assigns to another country j in year t. Typically, these weights sum to one (known as row standardization). The choice of weight is a crucial one because it determines the responsiveness of i to j's policy (given by $\rho\omega_{i,j,t}$). A number of different types weighting schemes have been used in the literature. The simplest of these is an equal weighting scheme, i.e. with N countries, $\omega_{j,i,t} = 1/N$.

If it can empirically be shown that other countries' tax policy is causal for a country's tax policy choices, then the fourth alternative outlined above – changes in common underlying factors – would become less plausible. In a properly specified econometric model, a common change in, say, the ideologic shift towards tax policy, would be captured by the control variables, not by the spatial lag coefficient. Thus, these studies imply that the observed relationship between taxes in one country and those elsewhere is not due such a common shift. However, they do not allow for a distinction between the other explanation approaches..

Regularity #2b: Countries react to their neighbours' tax rate changes within regions (clustering). A frequently used class of weighting schemes is one based on geographic proximity. This can be a variant of the equal weights where only neighboring countries receive a positive weight (something used by Altshuler and Goodspeed (2007) and Devereux, Lockwood, and Redoano (2008) in their studies of corporate taxes). Alternatively, weights can be a declining function of the distance between countries i and j (as was employed by Davies and Naughton, forthcoming, in their study of environmental policies and by Davies and Vadlamannati, 2013, in their study of labor policy). Klemm and van Parys (2012) bring these two together and use inverse-distance weights within a region, but assume zero weights across regions. These distance schemes are based on the empirical regularity that FDI is deterred by distance between the parent and host (see Blonigen and Piger (2011) for a recent overview of FDI determinants). Thus, a tax rate reduction close to country i may be more suc-

[9] Note that this variable is often considered endogenous, either due to errors correlated across countries or due to "strategic" endogeneity as suggested by the competition literature. As discussed by Anselin (1988) and many others, there are several methods available for dealing with this which are not germane to our analysis. We do wish to point out, however, that one method for dealing with this endogeneity is to use $t-1$ values for the spatial lag.

cessful in attracting away i's capital, forcing it to respond more to a proximate country than a distant one.

In general, the significance of the estimated spatial lag coefficient increases when using these distance weighting schemes than when using equal weights. An implication of such a scheme is that if countries respond more to neighbours, this results in a geographic clustering pattern. This led Klemm and van Parys (2012), among others, to argue that tax competition for mobile capital is greatest within geographic areas. However, it must be noted that this does not preclude the possibility of yardstick competition driving such estimates. If the electorate of one country finds it easier to observe the policies set by proximate nations, this would result in a comparable pattern of tax reactions.

Regularity #2c: Large countries are leaders in international tax setting. Instead of distance-based weights, some studies such as Devereux, Lockwood, and Redoano (2008) use GDP-based weights, i.e. $\omega_{j,i,t} = GDP_{j,t} / \sum_{k \neq i} GDP_{k,t}$.[10] A second example of this is Davies and Vadlamannati (2013), who also use population-based weights.[11] The rational behind such weights is again driven by the FDI determinants literature which finds that FDI is generally attracted to large countries. This suggests that country i may be forced to respond more to tax cuts by large countries than small ones. Indeed, the theory of Haufler and Wooton (1999) shows that large countries may be expected to win over small ones when competing for FDI.

As with the distance weights, using size weights improves the significance of the spatial lag relative to equal weights. However, as there is no test for which weighting scheme is "correct", the empiricist must cannot definitively say whether size or distance matters more. What can be said, however, is that the estimates indicate that a given nation responds more to some nations – including large ones – than it does to others. Again, however, it must be noted that if it is easier to observe the policies of a large country, yardstick competition can again be driving the results.

An important distinction here is that the above discussion refers to large countries as "leaders" but that this does not imply leadership in the Stackelberg sense. There, leadership also conveys the ability to use one's own tax to manipulate the tax rate of others. In an unpublished working paper, Altshuler & Goodspeed (2007) consider whether the US has taken the position of

[10] Davies and Klasen (2013) also use GDP weights in their study of overseas development aid where the motivation was that when deciding the amount of aid to provide a recipient nation, a donor country gives greater consideration to the donations of large countries than small countries.

[11] Others use weighting schemes based on the "attractiveness" of a given location. Examples of this include Devereux, Lockwood, and Redoano's (2008) use of FDI as a weight, Exbrayat's (2009) phi-ness weights, and Davies and Voget's (2009) market potential weights. More recently, some papers have begun to employ multiple spatial lags to examine whether a given country responds differently to another based on some classification such as EU membership (Davies and Voget (2009), Redoano (2007)) or OECD membership (Davies and Naughton, 2006), or level of development (Davies and Vadlamannati, 2013).

a Stackelberg leader in international tax competition and find some supporting evidence.[12] Leadership within Europe by Germany or the United Kingdom cannot be confirmed, however. In the light of the other empirical studies discussed above, it is nevertheless evident that potential US leadership is only one part of the international tax setting story since both tax credit and tax exemption countries seem to react to one another.

Regularity #3: More open economies react more strongly. In addition to considering whether the response of given nation i to the tax of j depends on j's characteristics, some evidence documents an effect of i's characteristics on its resonse. For example, Devereux et al. (2008) report that countries with less strict capital controls are more sensitive to their neighbours' tax policy choices. This could effectively be a strong hint to tax competition as the main cause for spatial correlation would be associated with capital flows between countries.[13] However, more open countries may also be those where it is easiest to observe policies set elsewhere, bolstering the case for yardstick competition. Furthermore, Overesch & Rincke (2011) reexamine this linkage and show that it is not robust. Thus, while there is some suggestion that more open economies react more strongly to tax changes elsewhere, this must be treated with some caution.

Regularity #4:: Stickiness of corporate tax rates. A characteristic feature of international tax rate setting is the infrequent rate of adjustment. Heinemann et al. (2010) report that, on average, national governments change their tax rate every four years. When combined with the typical legislative period's four year length, this means that the average elected government changes the tax rate once. This hints at some association with the political system, and is thus suggestive of yardstick competition. However, it is equally plausible that it stems from the new government re-assessing the marginal cost of public funds or deriving new estimates of the capital demand elasticities and then adjusting the tax rate. Alternatively, this could also stem from an adjustment cost indicating that it is only when the current tax policy becomes sufficiently divergent from current conditions that a government makes an adjustment. As such, Heinemann et al.'s results indicate that a tax rate adjustment cost is a desirable feature of a theory of tax setting.

Towards a learning based theory of tax competition In the next section, we present a model of a learning based model of tax policy diffusion. We

[12] Stackelberg leadership by the large(st) country is usually simply assumed as in Gordon (1992) or Baldwin & Krugman (2004). It should be noted that Kempf & Rota-Graziosi (2010) show that, if one endogenizes the leadership role, it is likely that the small country leads the tax competition.

[13] This is not to say that there are no studies which make the plausible case for yardstick competition on the community level. For instance, Bordignon et al. (2003) show for Italian municipalities that interaction in tax rates only occur in jurisdictions with high degrees of political competition.

will argue that the above discussed findings and data pattern may be explained by countries learning from each other without any tax competition or yardstick competition involved. In doing so, our theory will explain the persistent downward trend in tax rates, the seemingly causal relationship between tax rate cuts and the stickiness of corporate tax rates. Nevertheless, our model differs in critical ways from the explanations discussed above. First, in our model there are no real externalities in the form of capital flows or anything else. Because of this, a change in one country's policy does not impact the payoff to the policies chosen by another country. Nevertheless, unlike capital, information does move between countries, creating an information externality. Second, our approach differs from yardstick competition because it does not rely on an asymmetric information problem (but rather on an incomplete information problem). Therefore the information spillovers affect choices in a very different manner than in yardstick competition. Moreover, there are no political economy aspects to our story (although they would be straightforward to integrate). Third, the shared tax movements in our model come about not due to changes in a common underlying parameter, but due to the arrival of new information, both from a country's own experience and that of others.

With that said, we see our approach as complementary to the explanations discussed above.[14] In no way do we seek to suggest that countries do not compete for mobile capital or that electorates do not base their voting decisions in part on their observations of what transpires elsewhere. While we could embed our learning-based model in a tax competition or yardstick competition model, we instead shut down those channels of policy diffusion in order to focus on the specific contributions our approach makes.

3 Model

Intuition We consider a model with many countries, the welfare of which depends on a non-observable state of nature. For each potential state of nature, there is an adequate policy response. By observing their past policy experience, countries form beliefs on the true state of nature and adjust their policies accordingly. If the unknown state of nature is common to all countries within a group (a network), a country may not only use its own experience but may infer the other countries' experience by looking at their policy choices. If the inference is correct, the behavior will yield conditional convergence - something which is frequently observed in the data.

As an example, consider the unknown state of nature to be the extent of tax evasion (which is, by definition, not easy to observe) or, more generally, the ex-

[14] Brueckner (2003) offers a taxonomy of models of strategic government interaction. He proposes two categories, spill-over models and resource-shift models. In spill-over models, the objective function of an individual government directly depends on the actions taken by other governments. In resource-shift models, the objective functions depend on the quantity of resources within its borders which itself depends on the actions taken by all governments. Both models give rise to reaction functions. The learning based theory of international tax rate setting is different from both types of models.

tent to which agents react to taxation. Simply by observing current tax revenue, a country's government cannot precisely infer the degree of taxpayers' elasticity towards taxation. Inference requires a sufficiently large number of observations to identify the true relationship between taxes and payoffs. Observing other countries' tax policy choices increases the number of observations and, thus, the preciseness of the elasticity estimates.

An important feature of our model is that, in contrast to the tax competition or yardstick competition models, the welfare of an individual country does not depend on the tax rates set elsewhere. Similarly, a country's own tax policy will affect its own tax base and has no effect on other countries' welfare. We will show that, nevertheless, tax policy may be inefficient because a tax rate change reveals information and, thus, makes other countries' belief formation more efficient.

Model setup Consider a setting with $N \geq 2$ infinitely-lived countries which are indexed by i. Countries are ex-ante characterized by a vector of local, time-invariant characteristics $\mathbf{x}_i$ and an initial policy scalar $T_{i,0}$ which we will refer to as the "tax rate" for simplicity of exposition.[15] In addition, there is a common state of nature S which is drawn in period 0 from a cumulated distribution function $F_S(S)$ where $f_S(S) = F_S'(S)$ denotes density function. The true state of nature is unobserved by countries. In each period, the government in country i sets a tax rate $T_{i,t}$. This happens simultaneously across countries. If $T_{i,t} \neq T_{i,t-1}$, the government incurs a one-time adjustment cost $\alpha \geq 0$. Following tax setting, a gross-of-adjustment cost payoff, $w_{i,t}$, is received, given by

$$w_{i,t} = U(T_{i,t}; \mathbf{x}_i, S) + \varepsilon_{i,t} \tag{2}$$

where $\varepsilon_{i,t}$ is a random variable distributed independently around its mean zero. The randomness prevents i from perfectly extrapolating S from its payoff. Given $T_{i,t}$

We assume that country i can observe past tax rate choices of a set of countries denoted by Ω_i. Following Gale and Kariv (2003), this would define the countries that i is connected to, thereby defining the information network architecture. [16] Furthermore, $F(.)$, $H(.)$ and the $\mathbf{x}_i$s are common knowledge within the network Ω_i. Whereas the $T_{i,t}$ can be observed as soon as they are set for $i \in \Omega_i$, the $w_{i,t}$ is private information of country i.[17] Thus, in period t, country i has an information set $\mathbf{I}_{i,t} = \left\{ \{w_{i,k}\}_{k=0}^{t-1}, \{T_{j,k}\}_{k=0}^{t-1} \right\}$ for $j \in \Omega_i$ consisting of its own payoff history and the tax rates by other countries in the network.

Given this information, country i forms a belief about the true state of

[15] We could allow $\mathbf{x}_i$ to vary over time, however, as long as it is known in each period, implying that the government is able to use its payoff to make inferences on the state of nature, none of our results would change.

[16] For the moment, we treat the set Ω_i as constant over time.

[17] The case in which $w_{i,t}$ is observable by other countries will be discussed later on.

nature, denoted by

$$b_{i,t} = E\left(S|\mathbf{I}_{i,t}\right) = \int_{-\infty}^{+\infty} \tilde{S} f_S\left(\tilde{S}|\mathbf{I}_{i,t}\right) d\tilde{S} \tag{3}$$

$f_S\left(\tilde{S}|\mathbf{I}_{i,t}\right)$ denotes the conditional density of an arbitrary value of S denoted by $\tilde{S}$ given the information set $\mathbf{I}_{i,t}$.

Let $\hat{T}_{i,t} \equiv \arg\max_{T_i}\left\{w_i(T_i)\,|\hat{b}_{i,t}\right\}$ be the tax rate that maximizes the current payoffs for given beliefs $\hat{b}_{i,t}$. The arguments $\mathbf{x}_i$ and S in the payoff function will be suppressed for presentational ease. We will refer to $\hat{T}_{i,t}$ as the information-conditioned optimal (ICO) tax.[18] From the viewpoint of period t, $\hat{T}_{i,t+1}$ is a stochastic variable with distribution function $G_t(.)$ with mean $\hat{T}_{i,t}$, i.e. $E_{\hat{b}_{i,t}}\hat{T}_{i,t+k} = \hat{T}_{i,t}$ for all k, and variance $\sigma_t^{\hat{T}}$, $G_t\left(\hat{T}_{i,t}, \sigma_t^{\hat{T}}\right)$. The variance $\sigma_t^{\hat{T}}$ declines in t. Note that when $\Omega_i \neq \emptyset$, the beliefs of i in t depends on the tax rates elsewhere in $t-1$ and, as a result, so too will its tax rate in t. This has a parallel in the spatial lag literature where the weighted sum of $t-1$ taxes elsewhere are used as an explanatory variable for that tax rates a given country sets in t. As a final piece of notation, denote the optimal tax, that is, the one that would maximize expected welfare if the true state of nature were known with $T_i^*\left(\mathbf{x}_i, S\right)$.

Belief formation, policy choice and learning in isolation As a baseline, consider the case in which $\Omega_i = \emptyset$ for all i, that is, countries do not observe one another's choices, i.e. the model is described by N unconnected networks.[19] This implies that country i's information set, $\mathbf{I}_{i,t}$, has $2t$ elements, and is given by $\mathbf{I}_{i,t} = \left\{\{T_{i,k}, w_{i,k}\}_{k=0}^{t-1},\right\}$.

In isolation, the conditional density of an arbitrary value of $\tilde{S}$ is calculated, using Bayes' formula, by

$$f_S\left(\tilde{S}|\mathbf{I}_{i,t}\right) = \frac{\prod_{s=0}^{t-1} f_w\left(w_{i,s}|\left(\tilde{S}, T_{i,s}\right)\right)\cdot f_S\left(\tilde{S}\right)}{\prod_{s=0}^{t-1} f_w\left(w_{i,s}|T_{i,s}\right)} \tag{4}$$

where $\prod_{s=0}^{t-1} f_w\left(w_{i,s}|T_{i,s}\right) = \int_{-\infty}^{+\infty}\left[\prod_{s=0}^{t-1} f_w\left(w_{i,s}|\left(\tilde{S}, T_{i,s}\right)\right)\cdot f_S\left(\tilde{S}\right)\right] d\tilde{S}$.

The following lemma characterizes the belief formation under isolation.

Lemma 1 *(i)* $lim_{t\to\infty} b_{i,t} = S$.
(ii) $\mathrm{E}(b_{i,t+s}|\mathbf{I}_{i,t}) = b_{i,t}$ *for all* s.

[18] Although it is not necessary, to ease exposition, we will often treat this as single-valued for all x_i and $\hat{b}_{i,t}$.

[19] This is equivalent to a setting in which the state of nature is country specific and is independently determined across countries, something discussed below.

Proof. To be added. ∎

In the absence of adjustment cost, $\alpha = 0$, any change in beliefs will trigger a change in the tax rate $T_{i,t}$. Put differently, in each period t, $T_{i,t} = \hat{T}_{i,t}$. If there is a non-zero adjustment cost, the tax rate will only be adjusted if the discounted stream of surplus is at least equal to the adjustment cost. The following lemma proves that it suffices to compare the expected welfare from sticking to the tax rate in place forever to the welfare from changing it once and then sticking to it forever.

Lemma 2 *Country i changes the tax rate to $\hat{T}_{i,t}$ whenever*

$$\frac{\int_{-\infty}^{+\infty}\left[U\left(\hat{T}_{i,t};\mathbf{x}_i,\tilde{S}\right)-U\left(T_{i,t-1};\mathbf{x}_i,\tilde{S}\right)\right]\cdot f_S\left(\tilde{S}|\mathbf{I}_{i,t}\right)d\tilde{S}}{1-\delta}\geq\alpha \tag{5}$$

Proof. To be added. ∎

We can now state the following proposition.

Proposition 1 (Learning in isolation) *Assume that $\Omega_i = \emptyset$. (i) For $t \to \infty$, learning is perfect and, given the amount of information, efficiently quick.*
(ii) Assume that α is not prohibitively large. There is conditional convergence in tax rates. In the presence of adjustment costs, convergence remains imperfect.

Proof. (i) Following Lemma 1, $b_{i,t}$ converges to S as time passes on. In this sense, learning is perfect. Since learning does not depend on tax rate adjustments (and, thus, not on the adjustment cost) learning is efficiently quick. (ii) Since S is common for all countries, there is convergence, conditional on $\mathbf{x}_i$. Lemma 2 proves that tax rates may not be adjusted to their final value $T^*(S)$. ∎

If all countries are learning in isolation but share a common S, tax rates will tend to converge simply as a consequence of the converging beliefs. It is notable that this is precisely what is found in Slemrod (2004) who reports that corporate tax rates displayed significant convergence in tax rates from 1975 to 1995.

Belief formation, policy choice and learning in completely connected networks Now, consider the case of $\Omega_i \neq \emptyset$. Countries in Ω_i share a common S and observe all others' tax choices. In the terminology of Gale and Kariv (2003), this is a completely connected network.[20] By choosing $T_{i,t}$, country i conveys information to observers about its beliefs and, thus, about S. Thus, country i's belief about S in period t depends not just on its own experience, but on what it is able to intuit from the actions of others. As such, its information set $\mathbf{I}_{i,t}$ is substantially larger than when learning is in isolation. Instead of having $2t$ elements, the information set now has $(N+1)\,t$ elements, i.e. $\mathbf{I}_{i,t} = \left\{\{T_{i,k}, w_{i,k}\}_{k=0}^{t-1}, \{T_{j,k}\}_{k=0}^{t-1}\right\}$ for all $j \in \Omega_i$.

[20] We consider the case in which countries observe a subset of others, i.e. an incomplete network, below.

We start by considering the case of costless adjustment. Then, as in isolation, each country i will set $T_{i,t} = \hat{T}_{i,t}$ for all t. Therefore, by observing country j's tax rate, $\hat{T}_{j,t-1}$ with $j \in \Omega_i$, country i can perfectly induce j's welfare level $\tilde{w}_{j,t-2}$ and beliefs $\tilde{b}_{i,t}$ (where the tilde represents inferred values).[21] Since in equilibrium all countries use the information efficiently, i.e. correctly, and identically, the effective information set, denoted by $\hat{\mathbf{I}}_{i,t}$, is given by $\hat{\mathbf{I}}_{i,t} = \left\{ \{T_{i,k}, w_{i,k}\}_{k=0}^{t-1}, \{T_{j\neq i,k}, \tilde{w}_{j,k}\}_{k=0}^{t-2}, \{T_{j\neq i,t-1}\} \right\}$ for all $j \in \Omega_i$. Thus, the information sets across countries are equal with the exception of $w_{i,t-1}$ which is found in $\hat{\mathbf{I}}_{i,t}$ but not in $\hat{\mathbf{I}}_{-i,t}$.[22]

When adjustment is costly, some countries may not adjust their tax rates in each period because the expected gain from adjustment does not outweigh the expected cost. In these cases, information on payoffs and beliefs is only imperfectly transmitted. The other countries can, nevertheless, deduce information from the fact that the tax rate has not been adjusted. Each time a country does not adjust, there is a loss in information. If the country adjusts in a later period, some of the information is retrieved, but only imperfectly. Note that the estimation errors do not net out across different country observations since deviations in both directions cause an expected loss in welfare. Let $\tilde{w}^e_{j\neq i,k}$ denote the estimated welfare of other countries (from the viewpoint of country i). Then, country i's information set is given by $\hat{\mathbf{I}}^e_{i,t} = \left\{ \{T_{i,k}, w_{i,k}\}_{k=0}^{t-1}, \left\{T_{j\neq i,k}, \tilde{w}^e_{j,k}\right\}_{k=0}^{t-2}, \{T_{j\neq i,t-1}\} \right\}$ for all $j \in \Omega_i$.

In completely connected networks, the conditional density of an arbitrary value of $\tilde{S}$ is calculated, using Bayes' formula, by

$$f_S\left(\tilde{S}|\hat{\mathbf{I}}^e_{i,t}\right) = \frac{\prod_{s=0}^{t-1} f_w\left(w_{i,s}|\left(\tilde{S}, T_{i,s}\right)\right) \cdot \prod_{j\in\Omega_i} \prod_{s=0}^{t-2} f_w\left(w_{j,s}|\left(\tilde{S}, T_{j,s}\right)\right) \cdot f_S\left(\tilde{S}\right)}{\prod_{s=0}^{t-1} f_w\left(w_{i,s}|T_{i,s}\right) \cdot \prod_{j\in\Omega_i} \prod_{s=0}^{t-2} f_w\left(w_{j,s}|T_{j,s}\right)} \tag{6}$$

where $\prod_{s=0}^{t-2} f_w\left(w_{j,s}|T_{j,s}\right) = \int_{-\infty}^{+\infty} \left[\prod_{s=0}^{t-2} f_w\left(w_{j,s}|\left(\tilde{S}, T_{j,s}\right)\right) \cdot f_S\left(\tilde{S}\right)\right] d\tilde{S}$.

The following lemma characterizes the belief formation under isolation.

Lemma 3 *In completely connected networks, Lemma 1 holds (with* $\mathbf{I}_{i,t}$ *being replaced by* $\hat{\mathbf{I}}^e_{i,t}$*).*

Proof. To be added. ■

The subsequent lemma describes the tax setting in completely connected networks.

[21] Note that this type of learning does not begin until $t = 2$ since the welfare levels before period 0 do not convey information on S (which is drawn at the beginning of period 0).

[22] Note that the effort to reach these inferred beliefs can be substantial as they are all interrelated. We discuss the idea of rational inattention, one form of limited processing ability, below.

Lemma 4 *Country i changes the tax rate to $\hat{T}_{i,t}$ whenever*

$$\frac{\int_{-\infty}^{+\infty}\left[U\left(\hat{T}_{i,t};\mathbf{x}_i,\tilde{S}\right)-U\left(T_{i,t-1};\mathbf{x}_i,\tilde{S}\right)\right]\cdot f_S\left(\tilde{S}|\hat{\mathbf{I}}^e_{i,t}\right)d\tilde{S}}{1-\delta}\geq\alpha \tag{7}$$

Proof. To be added. ■

Since $\hat{\mathbf{I}}^e_{i,t}$ is much larger than $\mathbf{I}_{i,t}$ in isolation, policy choices can be based on more and, in expected terms, more precise information.[23] Tax rate adjustments in country $j \in \Omega_i$ affect the information set $\hat{\mathbf{I}}^e_{i,t}$. This is characterized by the following Lemma.

Lemma 5 *A tax rate adjustment in country j increases expected welfare in country $i \neq j$.*

Proof. For the proof, we compare the situations with and without a tax rate adjustment in j. Assume that country i is on the edge of being indifferent between changing and keeping the tax rate $T_{i,t-1}$. Without a change in j's tax rate, country i will stick to $T_{i,t-1}$. Then, a change in $T_{j,t}$ reveals information such that either i keeps the tax rate or it changes it to $\hat{T}_{i,t}$. In the former case, the tax rate change does not affect country i, in the latter country i is either indifferent or better off. That is, country i benefits in expected terms from more information due to a tax rate change in j. ■

Lemma 5 implies that beliefs become more accurate and $\left|\hat{b}_{i,t}-S\right|$ smaller the larger the information set.

We can now state Proposition 2.

Proposition 2 *Assume that $\Omega_i \neq \emptyset$ and α is not prohibitively large. (i) Learning (and convergence) are faster than in isolation as long as α is not too large. (ii) The speed of learning (and convergence) increases in N and decreases in α. (iii) If $\alpha > 0$, the number of tax rate changes (and, thus, the speed of learning) is inefficiently low.*

Proof. Parts (i) and (ii) are a direct consequence of the larger size of the effective information set in completely connected networks. If α is too large, no country will ever adjust and, as a consequence, inference becomes impossible. The effective information set is then equal to the one under isolated learning. Part (iii) follows directly from Lemma 5. ■

Part (iii) of the above Proposition points to a new type of externality in tax setting. It arises because adjustment carries greater information to other countries, allowing them to refine their beliefs and make better choices on whether to adjust themselves and, if so, what tax to use. In principle, the countries could coordinate themselves and subsidize each other's tax rate adjustment. Note, though, that there is no clear direction of the bias in tax rates that will occur due to the externality. Learning is inefficiently slow, but the theory does not

[23] Note that this is in expectation because through happenstance the information can be inaccurate, leading country i to temporarily move its beliefs away from the true state of nature.

allow to derive whether optimal tax rates are approached from above or from below something that will depend both on the initial tax rates $T_{i,0}$ and on the random payoff components in the early periods (where they have a large effect on beliefs).

Implications for empirical testing In the following, we describe and discuss a number of implications for empirical research.

First, whereas convergence in tax rates occurs under isolated learning and in complete networks, our model predicts that countries react to each others' tax rate settings if they are in the same network. In contrast to the existing theories, the co-movement of tax rates does not rely on the tax rate in one location changing the payoffs in another. The externality is a pure information externality. Empirical findings that allow for establishing causal relationships between the neighbours' tax rates and a country's own tax rate setting are thus in line with our model.

Second, the tax rate pattern predicted by our model looks as if tax rates were strategic complements, i.e. an increase in a tax rate in one country may trigger an increase in a tax rate elsewhere (precisely the finding in the empirical tax competition literature, e.g. Devereux, Lockwood, and Redano, 2008). For the sake of exposition, assume that $\mathbf{x}_i$ is the same across countries. Then,countries for which $T_{i,t-1} \neq \frac{1}{N-1} \sum_{k \neq i} T_{k,t-1}$ will on average revise their beliefs so that the ICO tax is the average tax rate of all countries.[24] This implies that lower than average tax countries will increase their tax rates while higher than average countries will cut theirs.

Third, the speed of convergence increases in the number of network members. While it may seem difficult to identify networks in the data, it is rather common than the exception to assume some sort of network in the existing studies which usually presume that, for instance, tax competition within the EU is different from tax competition with the US. An interesting effect occurs when the network becomes larger, for instance, if the EU is enlarged (see Davies and Voget, 2009).

Fourth, if adjustment is costly, the countries most likely to adjust their tax rates will be those whose tax rates differ the most from this average – a finding in Heinemann et al. (2010). This is because countries who are near the world average will find that the information which comes in from others does not impact their beliefs by enough to induce action.

A fifth – and somewhat speculative – implication of the learning process is that some countries may feature more heavily in forming the beliefs of others. In particular, a country with less noise in its own payoff, will on average have more accurate beliefs on the true state of nature after the first period.[25] Because the information content in such a country's actions are greater, other nations will give greater weight to that nation's actions when revising its beliefs. This would then give a new justification for using non-equal weights in a spatial lag

[24] Note that this is on average because the period t beliefs will still include information on $w_{i,t}$ (meaning that information sets still differ somewhat).

[25] In fact, when $\sigma_i^{\varepsilon} = 0$, i would be able to perfectly deduce S after one period.

regression. For example, if the payoff noise is inversely related to country size, this would motivate the use of GDP or population weights.[26]

Incompletely connected networks Now assume that each country only observes an exogenous subset of the other countries, i.e. where Ω_i is smaller than the set of countries for at least some i. This is known as an incompletely connected network (Gale and Kariv, 2003). This network structure can result from either a limited flow of information (such as when a country observes only the tax rates of its neighbors) or a limited ability to process the incoming information (such as the rational inattention literature of Sims (1998, 2003) and others).[27]

In the extreme, the set of countries is simply divided into a collection of separate, non-overlapping groups where either states of nature are correlated within groups but not across groups or in which countries observe in-group actions, but not out-of-group actions. Here, the situation is simply a repeated version of the one discussed above with the prediction that what transpires within one peer group has no impact on what happens elsewhere. Indeed, Klemm and van Parys (2012), who estimate responsiveness of the tax incentives of one nation to those elsewhere find evidence of diffusion within Latin America and Africa, but not across regions. Thus, if geography inhibits the flow of information (or separate states of nature are drawn for different regions), then our model could result in such a pattern of tax policy responsiveness. These separate groups could also arise from other factors such as the level of development or ideology that inhibit the ability of a member of one group to collect or interpret the information from a member of another group.

Alternatively suppose that the Ω_is intersect with one another, i.e. the network is incompletely connected. As discussed in Gale and Kariv (2003), there are a number of different network architectures that can be considered, however

[26] One way to extend the model in this fashion is to let each nation be the aggregation of a number of local jurisdictions, each of which receives their own independently-drawn payoff comparable to $w_{i,t}$. As the number of jurisdictions increases, this would represent an increase in the number of observations flowing to the national government each period, improving the accuracy of its beliefs just beliefs become more accurate over time. Alternatively, a case can be made that the information gathered by developed nation governments may be of a better quality than what is collected by the government of a developing country. As such, it would be appropriate to give greater weight to the actions of developed countries both when forming beliefs and in constructing a weighting scheme for spatial estimation. Davies and Naughton (forthcoming) find that countries are more sensitive to the environmental policy of OECD members than to those of non-members. A similar difference in the quality of information could be linked to ideology or culture with the quality of information relatively noisy in nations where individual privacy is a primary concern. Although we are unaware of any study that has attempted to do so, one method of constructing such a weighting scheme for estimation would be to use the Hofstede (2001) "Individualism versus Collectivism" measure. The Hofstede cultural measures were found to be significant predictors of FDI by Davies, et. al (2008).

[27] One example of the costs of processing information and its impact on tax policy is found in information exchange, the process by which one tax authority provides information to another regarding a multinational active in both. Although intended to curb tax evasion, the cost of collecting and translating information can be a significant barrier. Vann (1996) and Krabbe (1996) provide additional discussion.

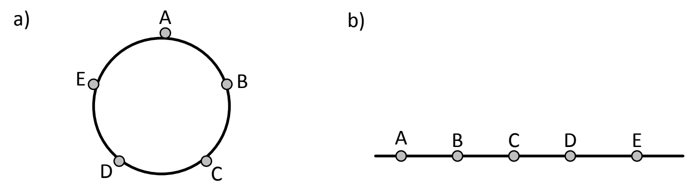

Figure 2: Incompletely connected networks

all converge (conditional on the $\mathbf{x}_i$s) to the same equilibrium actions in a finite amount of time. The rate of this convergence, however, does depend somewhat on the network architecture, that is, who observes the actions of who.

Consider, for instance, the network displayed on the left hand side in Figure 2. Here, the N countries are distributed on a circle and observe only the actions of their neighbors. More concretely, let France (A) observe Germany (B) and Spain (E), Germany (B) observe France (A) and Austria (C), and so forth. Starting in period 1, France's tax rate is based solely on its own experience, $(w_{A,0}, T_{A,0})$, and the a priori distribution of S. In period 2, France has observed the first period tax rates in Germany and Spain and has inferred their experience, $\left(\tilde{w}^e_{E,0}, T_{E,0}\right)$ and $\left(\tilde{w}^e_{B,0}, T_{B,0}\right)$. In addition, it has its own experience history, $(w_{A,0}, T_{A,0})$ and $(w_{A,1}, T_{A,1})$. In period 3, France again observes tax rates in Germany and Spain, but knows now that they are based on information from these two countries as well their direct neighbours. Of course, France is unable to correctly induce the welfare levels $w_{i,1}$ from one observed tax rate, but it may learn about the average experience in these countries that led to the choice of $T_{i,1}$. In period 4, information from two more countries enter the information set of France via their influence on France's neighbors' neighbors, and so forth. In general, an individual country learns about its direct neighbours' experience with a lag of one period, its second direct neighbours with a lag of two periods, its third direct neighbours with a lag of three periods, and so on. Thus, learning is efficient, albeit in a more limited way (since a country does not directly observe non-neighbor taxes) and with a longer delay. An implication of the network structure is that an individual country i will give larger weights to the observations from its neighbours than to its own experience, with the weights increasing over time. This is because neighbours' actions are based on i's own tax history (which i already knows) as well the tax history of nations that i cannot directly observe.

From the perspective of the empirical literature, incompletely connected networks would lead to spatial lags that include only the policies of neighboring countries, as done by Altshuler and Goodspeed (2007). A more general model would be based on the assumption that a country observes the actions of others, but that observation is distorted or delayed with the noise or lag increasing with distance. In that setting, although distant countries would still influence the

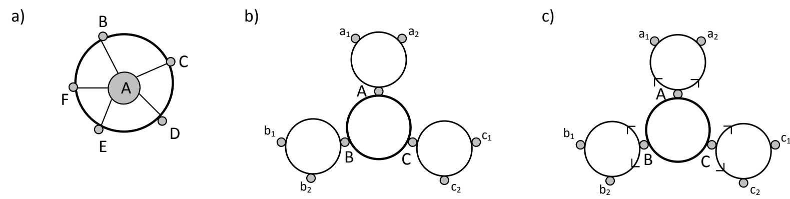

Figure 3: Hubs

beliefs and taxes of a given country, their importance would be muted. As such, the spatial lag could be constructed with weights that are decreasing in distance, something done by Davies and Vadlamannati (2013) and others. In particular, Davies and Naughton (forthcoming) discuss the different implications of various distance-based weighting scheme in their study of environmental policy diffusion.

An alternative network structure is depicted on the right hand side of Figure 2. While the circle on the left hand side implies a symmetry across countries, the line in b) implies that there are dead ends ("peripheries") on the left and on the right.[28] After period 2, countries B and D will give more weight to observations from C than from A or E since they contain more information. Similarly, countries A and E will give more weight to observations from B and D relative to their own experience since they contain more information. In such a setting, country C ("the center") has, from period 3 onwards, the largest information set and, thus, the most precise belief. In an empirical setting, it would therefore have a greater than average impact on the taxes of others.

A twist on the incomplete network structure is one with overlapping networks. Figure 3 presents examples which build from the circular structure of Figure 2, panel a. In Figure 3a, A is a large country surrounded by smaller countries B to F. Whereas A can observe each other country's actions, the smaller ones only observe their direct neighbours including country A. Again, there will be different learning speeds and differently accurate beliefs. Country A has the most and best (i.e. most recent) information. Therefore, each of the small countries will base its beliefs more heavily on the large country's past actions than on those of other small countries. In this sense, the hub nation A becomes acts as a "leader" and has a greater impact on the policies of other countries. Altshuler and Goodspeed (2007) find that other OECD countries' corporate taxes respond more to the US tax rate than to that of neighboring countries. This would be consistent with the type of structure in Figure 3a with the US acting as a hub.

In Figure 3b, a multi-centric world is depicted with three large countries (A, B and C) each of which has its own "backyard" or subnetwork (a_1, a_2 and so on). Whereas the large countries observe each others' actions and their subnetwork,

[28] Baldwin and Krugman (2004) provide a model of tax competition between a core country and the periphery.

the smaller backyard countries only observe each other and the large country in their subnetwork. Such a setting could occur when observation capacity is linked to trade flows. Whereas small countries only trade with the largest country in the region, the large countries engage in interregional trade. Now, there are three 'leaders' with superior information that get larger learning weights from their backyard countries. From an empirical point of view, spatial lags could be weighted with trade flows in order to capture connectedness.[29]

Regardless of the precise structure of the network, however, three results carry over from the completely connected model. First, because even in an incompletely connected network more information arrives to each i than in isolation, learning in the group is faster than in isolation. However, this learning is slower than in a completely connected network.[30] Second, although gradual, learning in an incompletely connected network is perfect in a finite amount of time. Third, in the presence of positive adjustment costs, there is underadjustment relative to the social optimum due to the uninternalized externality of conveying information. An important difference to the case of completely connected networks is, however, that the speed of learning may differ across countries.

Proposition 3 *Prop. 2 holds in incompletely connected networks if* $\Omega_i \notin \emptyset$ *for all* i.

Proof. The proof follows from the positive expected value of an additional piece of information. ■

Non-shared states of nature In the above, we assumed that all countries in a network shared the same state of nature. Alternatively, assume that the states of nature S are not identical across countries, but are correlated with one another, though imperfectly.[31] The S^i are assumed to be jointly distributed according to some joint cumulated distribution function $F(S_1, ..., S_N)$. Observing the behavior of other countries still conveys useful information but due to the imperfect correlation of S_i across countries, the informational content is limited. Nevertheless, as $t \to \infty$, learning remains perfect and each country's tax will converge to $T_i^*(\mathbf{x}_i, S_i)$ (up to the limits created by adjustment costs). The speed of this convergence will lie between that when learning in isolation and when S is common across countries. As such, there are still welfare gains relative to learning in isolation and an unrealized positive externality from adjustment.

Another implication is that countries will extract more information from countries whose states of nature are more correlated with their own. For example, states of nature may be more correlated between proximate countries than distant ones. This would provide a rationale for using a distance-based weighting

[29] Although we do not know of such an estimation, Exbrayat (2009) uses a measure of overall trade openness as a weighting scheme.

[30] Gale and Kariv (2003) provide simulation results that compare the speed of convergence.

[31] If the country-specific states of nature are independently distributed, then no country learns anything of use by observing the actions of others. As such, the game is equivalent to one of learning in isolation.

matrix when estimating policy diffusion even in a completely connected network. Alternatively, the correlation between the S_is could be greater between those with more similar $\mathbf{x}_i$s, which could include factors such as ideology. This would suggest the use of political similarity weights, something utilized in Davies and Klasen (2013), who use a measure of political affinity based on United Nations voting as a weighting scheme for their study of overseas development assistance donations.

Taking this idea a step farther, one can imagine a setting in which a subset of nations, for instance the EU, share a common state of nature but others have individual states. In this case, although all countries learn from each other, there is a distinction between what is learned from an EU and a non-EU country.This would then fit the pattern found by Redoano (2007) and Davies and Voget (2009), who find that while non-EU members respond equally to the corporate taxes of both EU and non-EU countries, members respond less to the taxes of non-members than they do to members. In any case, the results of Proposition 3 holds, i.e. group learning increases welfare over the isolated learning case but happens inefficiently often in the presence of adjustment costs.

One-way information flows In the above described architectures, it is assumed that if two countries have a network link, information flows are bilateral, that is, if A observes B, B observes A as well. As an alternative, one can assume that for some links information flows move in only one direction. For example, we might assume that countries can only observe their two most important parters' actions which, in a multilateral world need not align across countries.[32] If importance is associated with trade, learning would unilateral for some links.[33] Part c) of Figure 3 illustrates a case in which the three large countries observe each others' actions but not those of smaller countries in their own subnetwork. The smaller countries, however, observe all of the actions in their subnetwork. Now, the small countries have the richest set of information as they learn about their own subnetworks' members' actions and, indirectly, about the two other large countries. In contrast, the large countries never glean information from small countries be they in or out of their subnetwork. As a consequence, the small countries' learning speed is higher and their beliefs are more accurate.[34] That notwithstanding, the results of Proposition 3 continue

[32]For example, the US and the UK are Ireland's primary trading partners, however, Ireland accounts for a much smaller share of trade for those two nations.

[33]Alternatively, the resources available to process information can vary across countries. If such resources are more limited for small countries than large ones, even if a large and a small country observe one another's actions, only the large country would be able to process the information, making the information flow effectively one-way

[34]A variant of this model could capture the relationship between the OECD (which is observed by all) and the non-OECD (who are only observed by one another). In this case, all nations would update their beliefs in response to OECD actions but OECD nations would not respond to the policies of non-OECD countries. This is the pattern found by Davies and Naughton (2006), who estimate the participation in international environmental agreements. Although they find that all countries respond to the participation of OECD countries, the same cannot be said of non-OECD countries' treaty participation decisions. In particular, for some specifications, although all countries respond to the OECD decisions, only non-OECD

to hold.

4 Conclusion

The two leading approaches to explain spatially correlated tax policies are competition for mobile resources (or: tax competition in the framework of tax policy) and yardstick competition. The goal of this paper has been to offer a third mechanism by which the policies chosen in one country can be influenced by those elsewhere. This mechanism operates via the information that is embodied in a country's policy choice. When the optimal policy depends on an unobserved state of nature which is correlated across nations, by observing the choices made in another location, a given country is able to refine its beliefs and base its decisions upon this. As a result, the policies of countries will converge on one another (up to the limits created by other differences across nations and potential adjustment costs). This then mirrors the patterns found in the empirical literature. A second implication, and one that mirrors the primary concern in the tax competition literature, is that the equilibrium pattern of adjustments can be inefficiently slow. This is because a given country does not internalize the benefits accruing to others from an adjustment in its own policy.

An important aspect of the model is that the policies of one country do not influence the mapping between policy and payoffs elsewhere. This is then in direct opposition to the mechanisms resulting in correlated policies in the tax competition or yardstick competition literatures. This is in no way intended to suggest that these mechanisms do not occur. Indeed, it is our supposition that all three sources of international policy diffusion jointly drive the patterns observed in the data. Instead, we hope to illuminate the additional interactions and the role that network architecture plays in those which result from social learning.

References

[1] Altshuler, R. & Goodspeed, T. (2007). Follow the Leader: Evidence on European and U.S. Tax Competition, Working Paper.

[2] Anselin, L. (1988). Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Boston, MA.

[3] Bala, V. & Goyal, S. (1998). Learning from Neighbours, Review of Economic Studies 65(3): 595-621.

[4] Baldwin, R. & Krugman, P. (2004). Agglomeration, integration and tax harmonisation, European Economic Review, 48(1), 1-23.

nations respond to non-OECD decisions in line with the predictions of this architecture.

[5] Beron, Kurt J., James C. Murdoch, and Wim P. M. Vijverberg (2003). Why Cooperate? Public Goods, Economic Power, and the Montreal Protocol, Review of Economics and Statistics, 85(2), 286-297.

[6] Besley, T. & Case, A. (1995). Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition, American Economic Review 85(1): 25-45.

[7] Blonigen, Bruce A. and Piger, Jeremy (2011). Determinants of Foreign Direct Investment. NBER Working Paper 16704.

[8] Bordignon, Massimo; Cerniglia, Floriana; Revelli, Federico (2003). In Search of Yardstick Competition: A Spatial Analysis of Italian Municipality Property Tax Setting, Journal of Urban Economics 54(2): 199-217.

[9] Brueckner, J.K. (2003). Strategic Interaction Among Governments: An Overview of Empirical Studies, International Regional Science Review 26(2): 175-188.

[10] Davies, R.B. (2005). State Tax Competition for Foreign Direct Investment: A Winnable War?, Journal of International Economics, 67(2), pp. 498-512.

[11] Davies, R.B., Ionascu, D. & Kristjánsdóttir, H. (2008). Estimating the Impact of Time Invariant Variables on FDI with Fixed Effects, Review of World Economy, 144(3), 381-407.

[12] Davies Ronald B. & Stephan Klasen (2013). Of Donor Coordination, Free-Riding, Darlings, and Orphans: The Dependence of Bilateral Aid on Other Bilateral Giving, CESifo Working Paper Series 4177, CESifo Group Munich.

[13] Davies, R.B. & Naughton, H.T. (forthcoming). Cooperation in Environmental Policy: A Spatial Approach, International Tax and Public Finance.

[14] Davies, R.B. & Vadlamannati, K.C. (2013). A Race to the Bottom in Labour Standards? An Empirical Investigation, Journal of Development Economics 103, 1-14.

[15] Davies, R.B. & Johannes Voget (2009). Tax Competition in an Expanding European Union, The Institute for International Integration Studies Discussion Paper Series No. 276.

[16] de Mooij R.A. & Sjef Ederveen (2008). Corporate tax elasticities: a reader's guide to empirical findings, Oxford Review of Economic Policy, Oxford University Press, vol. 24(4), 680-697.

[17] Devereux, M.P., B. Lockwood, & M. Redoano (2008). Do countries compete over corporate tax rates?, Journal of Public Economics 92(5–6): 1210–1235.

[18] Exbrayat, N. (2009). The impact of trade integration and agglomeration economies on tax interactions: evidence from OECD countries, Working paper.

[19] Gale, D. (1996). What Have We Learned from Social Learning, European Economic Review 40(3-5): 617-628.

[20] Gale, D. & Kariv, S. (2003). Bayesian Learning in Social Networks, Games and Economic Behavior 45(2): 329-46.

[21] Gordon, R., 1992. Can capital income taxes survive in open economies? Journal of Finance 47, 1159–1180.

[22] Heinemann, F.; Overesch, M. & Rincke, J. (2010). Rate-Cutting Tax Reforms and Corporate Tax Competition in Europe, Economics and Politics 22(3): 498-518.

[23] Heyndels, Bruno & Jef Vuchelen (1998). Tax Mimicking among Belgian Municipalities, National Tax Journal, 51: 89-101.

[24] Hofstede, G. (2001). Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations. New York: Sage.

[25] Krabbe, H. (1996). The Experience of Germany. In Richard Vann (ed.), Tax Treaties: Linkages Between OECD Member Countries and Dynamic Non-Member-Economies OECD.

[26] Levinson, Arik. (2003). Environmental Regulatory Competition: A Status Report and Some New Evidence, National Tax Journal, 56(1), 91-106.

[27] Markusen, J.R.; Morey, E.R. & Olewiler, N. (1995). Competition in Regional Environmental Policies When Plant Locations Are Endogenous, Journal of Public Economics 56(1): 55-77.

[28] Mintz, Jack & Michael Smart (2004). Income shifting, investment, and tax competition: theory and evidence from provincial taxation in Canada, Journal of Public Economics, 88: 1149-1168.

[29] Murdoch, James C., Todd Sandler, and Wim P. M. Vijverberg. (2003). The Participation Decisions versus the Level of Participation in an Environmental Treaty: A Spatial Probit Analysis, Journal of Public Economics, 87, 337-362.

[30] Olney, W.W. (2011). A Race to the Bottom? Employment Protection and Foreign Direct Investment. Williams College Economics Department Working Paper 2011-02.

[31] Overesch, M. and Rincke, J. (2009). Competition from low-wage countries and the decline of corporate tax rates: evidence from European integration, World Economy 32(9):1348–1364.

[32] Overesch, M. and Rincke, J. (2011). What Drives Corporate Tax Rates Down? A Reassessment of Globalization, Tax Competition, and Dynamic Adjustment to Shocks, Scandinavian Journal of Economics 113(3): 579-602.

[33] Redoano, M. (2007). Fiscal interaction among European countries. Does the EU matter?, CESifo Working Paper No. 1952.

[34] Salmon, P. (1987). Decentralization as an Incentive Scheme, Oxford Review of Economic Policy 3: 24–43.

[35] Slemrod, J. (2004). Are corporate tax rates, or countries, converging? Journal of Public Economics 88 1169-1186.

[36] Vann, R. (1996). Policy Dialogue with Dynamic Non-member Economies. In Richard Vann (ed.), Tax Treaties: Linkages Between OECD Member Countries and Dynamic Non-Member-Economies. OECD.

[37] Vives, X. (1996). Social Learning and Rational Expectations, European Economic Review 40(3-5): 589-601.

[38] Whalley, J., 1990. Foreign responses to US tax reform. In: Slemrod, J. (Ed.), Do Taxes Matter? The Impact of the Tax Reform Act of 1986. The MIT Press, Cambridge and London, pp. 286– 314.

[39] Wilson, J.D. (1986). A Theory of Interregional Tax Competition, Journal of Urban Economics 19(3): 296-315.

[40] Zodrow, G.R. & Mieszkowski, P. (1986). Pigou, Tiebout, Property Taxation, and the Underprovision of Local Public Goods, Journal of Urban Economics 19(3): 356-371.

Oxford University Centre for Business Taxation
Working Paper series recent papers

WP13/18 Reuven S Avi-Yonah *And yet it moves: taxation and labour mobility in the 21st century*

WP13/17 Anne Brockmeyer *The investment effect of taxation: evidence from a corporate tax kink*

WP13/16 Dominika Langenmayr and Rebecca Lesterz *Taxation and corporate risk-taking*

WP13/15 Martin Ruf and Alfons J Weichenrieder *CFC legislation, passive assets and the impact of the ECJ's Cadbury-Schweppes decision*

WP13/14 Annette Alstadsæter and Martin Jacob *The effect of awareness and incentives on tax evasion*

WP13/13 Jarkko Harju and Tuomos Matikka *The elasticity of taxable income and income-shifting between tax bases: what is "real" and what is not?*

WP13/12 Li Liu and Andrew Harper *Temporary increase in annual investment allowance*

WP13/11 Alan J Auderbach and Michael P Devererux *Consumption and cash-flow taxes in an international setting*

WP13/10 Andreas Haufler and Mohammed Mardan *Cross-border loss offset can fuel tax competition*

WP13/09 Ben Lockwood *How should financial intermediation services be taxed?*

WP13/08 Dominika Langenmayr, Andreas Haufler and Christian J bauer *Should tax policy favour high or low productivity firms?*

WP13/07 Theresa Lohse and Nadine Riedel *Do transfer pricing laws limit international income shifting? Evidence from European multinationals*

WP13/06 Ruud de Mooij and Jost Heckemeyer *Taxation and corporate debt: are banks any different?*

WP13/05 Rita de la Feria *EU VAT rate structure: towards unilateral convergence?*

WP13/04 Johannes Becker and Melaine Steinhoff *Conservative accounting yields excessive risk-taking - a note*

WP13/03 Michael P.Devereux, Clemens Fuest, and Ben Lockwood *The Taxation of Foreign Profits: a Unified View*

WP13/02 Giorgia Maffini *Corporate tax policy under the Labour government 1997-2010*

WP13/01 Christoph Ernst, Katharina Richter and Nadine Riedel *Corporate taxation and the quality of research & development*

WP12/29 Michael P Devereux and Simon Loretz *What do we know about corporate tax competition?*

WP12/28 Rita de la Feria and Richard Krever *Ending VAT Exemptions: Towards A Post-Modern VAT*

WP12/27 Theresa Lohse, Nadine Riedel and Christoph Spengel *The Increasing Importance of Transfer Pricing Regulations – a Worldwide Overview*

WP12/26 Harry Huizinga, Johannes Voget and Wolf Wagner
Capital gains taxation and the cost of capital: evidence from unanticipated cross-border transfers of tax bases

WP12/25 Harry Huizinga, Johannes Voget and Wolf Wagner
International taxation and cross border banking

WP12/24 Johannes Becker and Nadine riedel
Multinational Firms Mitigate Tax Competition

WP12/23 Michael Devereux, Li Liu and Simon Loretz
The Elasticity of Corporate Taxable Income: New Evidence from UK Tax Records

WP12/22 Olivier Bargain, Mathias Dolls, Clemens Fuest, Dirk Neumann, Andreas Peichl, Nico Pestel, Sebastian Siegloch
Fiscal Union in Europe? Redistributive and Stabilising Effects of a European Tax-Benefit System and Fiscal Equalisation Mechanism

WP12/21 Peter Egger, Christian Keuschnigg, Valeria Merlo and Georg Wamser *Corporate taxes and internal borrowing within multinational firms*

WP12/20 Jarkko Harju and Tuomos Kosonen *The impact of tax incentives on the economic activity of entrepreneurs*

WP12/19 Laura Kawano and Joel slemrod *The effects of tax rates and tax bases on corporate tax revenues: estimates with new measures of the corporate tax base*

WP12/18 Giacomo Rodano, Nicolas Serrano-Velarde and Emanuele Tarantino *Bankruptcy law and the cost of banking finance*

WP12/17 Xavier Boutin, Giacinta Cestone, Chiara Fumagalli, Giovanni Pica and Nicolas Serrano-Velarde *The Deep pocket effect of internal capital markets*

WP12/16 Clemens Fuest, Andreas Peichl and Sebastian Siegloch *Which workers bear the burden of corporate taxation and which firms can pass it on? Micro evidence from Germany*

WP12/15 Michael P. Devereux *Issues in the Design of Taxes on Corporate Profit*

WP12/14 Alan Auerbach and Michael P. Devereux *Consumption Taxes In An International Setting*

WP12/13 Wiji Arulampalam, Michael P. Devereux and Federica Liberini *Taxes and the location of targets*

WP12/12 Scott Dyreng, Bradley Lindsey and Jacob Thornock *Exploring the role Delaware plays as a tax haven*

WP12/11 Katarzyna Bilicka and Clemens Fuest *With which countries do tax havens share information?*

WP12/10 Giorgia Maffini *Territoriality, Worldwide Principle, and Competitiveness of Multinationals: A Firm-level Analysis of Tax Burdens*

WP12/09 Daniel Shaviro *The rising tax-electivity of US residency*

WP12/08 Edward D Kleinbard *Stateless Income*

WP12/07 Vilen Lipatov and Alfons Weichenrieder *Optimal income taxation with tax competition*

WP12/06 Kevin S Markle *A Comparison of the Tax-motivated Income Shifting of Multinationals in Territorial and Worldwide Countries*

WP12/05 Li Liu *Income Taxation and Business Incorporation: Evidence from the Early Twentieth Century*

WP12/04 Shafik Hebous and Vilen Lipatov *A Journey from a Corruption Port to a Tax Haven*

WP12/03 Neils Johannesen *Strategic line drawing between debt and equity*

WP12/02 Chongyang Chen, Zhonglan Dai, Douglas A. Shackelford and Harold H. Zhang, *Does Financial Constraint Affect Shareholder Taxes and the Cost of Equity Capital?*

WP12/01 Stephen R. Bond and Irem Guceri, *Trends in UK BERD after the Introduction of R&D Tax Credits*

WP11/24 Athiphat Muthitacharoen George R. Zodrow *Revisiting the Excise Tax Effects of the Property Tax*

WP11/23 Krautheim, Sebastian and Tim Schmidt-Eisenlohr *Wages and International Tax Competition*

WP11/22 Haufler, Andreas, Pehr-Johan Nörback and Lars Persson *Entrepreneurial innovation and taxation*

WP11/21 Mancini, Raffaele, Paolo M. Panteghini and Maria laura Parisi *Debt-Shifting in Europe*

WP11/20 Xing, Jing *Does tax structure affect economic growth? Empirical evidence from OECD countries*

WP11/19 Freedman, Judith *Responsive regulation, risk and rules: applying the theory to tax practice*

WP11/18 Devereux, Michael P. and Simon Loretz *How would EU corporate tax reform affect US investment in Europe?*

WP11/17 Vella, John, Clemens Fuest and Tim Schmidt-Eisenlohr *Response on EU proposal for a Financial Transaction Tax*

WP11/16 Loretz, Simon and Socrates Mokkas *Evidence for profit-shifting with tax sensitive capital stocks*

WP11/15 Weisenbach, david A. *Carbon taxation in the EU: Expanding EU carbon price*

WP11/14 Bauer, Christian, Davies, Ronald B. and Andreas Hauer *Economic Integration and the Optimal Corporate Tax Structure with Heterogeneous Firms*

WP11/13 Englisch, Joachim *National Measures to Counter Tax Avoidance under the Merger Directive*

WP11/12 de la Feria, Rita and Clemens Fuest *Closer to an Internal Market? The Economic Effects of EU Tax Jurisprudence*

WP11/11 Englisch, Joachim *EU Perspective on VAT Exemptions*

WP11/10 Riedel, Nadine and Hannah Schildberg-Hörisch *Asymmetric Obligations*

WP11/09 Böhm, Tobias and Nadine Riedel *On Selection into Public Civil Service*

WP11/08 Auerbach, Alan J. and Michael P. Devereux *Consumption and Cash-Flow Taxes in an International Setting*

WP11/07 Becker, Johannes and Clemens Fuest *Tax Competition: M&A versus Greenfield Investment*

WP11/06 Riedel, Nadine *Taxing Multinationals under Union Wage Bargaining*

WP11/05 Liu, Li and Rosanne Altshuler *Measuring the Burden of the Corporate Income Tax under Imperfect Competition*

WP11/04 Becker, Johannes and Clemens Fuest *The Taxation of Foreign Profits - The Old View, the New View, and a Pragmatic View*

WP11/03 Konrad, Kai *Search Costs and Corporate Income Tax Competition*

WP11/02 Hellerstein,Walter *Comparing the Treatment of Charities Under Value Added Taxes and Retail Sales Taxes*

WP11/01 Dharmapala, Dhammika and Nadine Riedel *Earnings Shocks and Tax-Motivated Income-Shifting: Evidence from European Multinationals*