The London School of Economics and Political Science

Spatial Dependence in Dyadic Data – The Cases of Double Taxation Treaties, Official Development Assistance, and Asylum Migration

Fabian Barthel

A thesis submitted to the Department of Geography and Environment of the London School of Economics for the degree of Doctor of Philosophy, London, October 2011

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Abstract

The thesis analyses spatial dependence in dyadic data by the means of three applications. These have in common that they concern bilateral international relations or flows between two countries with a particular focus on the relationship between developing and developed countries. While the first chapter provides a general introduction to spatial dependence with a focus on dyadic datasets, the second chapter looks at double taxation treaties (DTTs) and analyses whether strategic interaction among capital importing countries can explain the widespread conclusion of double taxation treaties between an industrialised and a developing country. This is important since upon entering such a treaty, the net-capital importer can lose a significant amount of tax revenues from foreign direct investment (FDI), while the net-capital exporter is better off. The analysis reveals that a country is more likely to enter a DTT if competitor countries for FDI also negotiated such a DTT, providing evidence for the hypothesis that the group of net-capital importers finds itself in a situation which can be described as a prisoners' dilemma: individually they would be better off if they refused to negotiate a treaty, but collectively they have an incentive to sign such a tax treaty. The third chapter is on official development assistance and deals with the question whether a specific donor tends to dedicate a larger share of its aid budget to a certain recipient if other donors give money to the same beneficiary. A considerable degree of spatial dependence is found in the form that donors tend to allocate their money to the same recipients. Donors particularly follow the example of the most important aid donors. This behaviour has negative implications for aid effectiveness, contributes to harmful aid volatility and leads to aid darlings and orphans. However, there is no evidence that donors strategically interact with each other in order to pursue their military strategic and economic goals. Spatial dependence in asylum migration is the third application, discussed in the fourth chapter. It is well documented in the literature that personal networks of migrants reduce the risk of migration and facilitate transition to the host country. So far it has always been assumed that these personal networks only exist for fellow countrymen. The empirical analysis, however, shows that the positive effects also operate across borders and that also migrants from other geographically close source countries make asylum migration from a given source country more likely. Furthermore, it is shown that a more restrictive asylum policy in one destination country provides a negative externality for other destinations. This is because asylum seekers are deflected by a tighter asylum regime and encouraged to lodge their application in more liberal target countries.

Acknowledgements

I would like to thank my family for all their support of any kind, particularly my parents for their apparently successful laissez-faire education and their steady encouragement to go my own path. I thank Viktoria for her support, loving inducement and for enduring a three-year long distance relationship. I further thank Oliver for irreplaceable assistance in finding solutions for my numerous STATA problems. A very special thanks to Mrs. Rosenauer without whom my life could easily have taken a less favourable direction.

My greatest thanks are to Eric Neumayer for the best supervision imaginable – both from a personal and a professional point of view. Without his commitment, great support and invaluable advice this dissertation would not have been possible – or at least much less fun. I am also grateful to Steve Gibbons for being my secondary supervisor and for many valuable comments on my first chapter. Last but not least I acknowledge financial support from the Department of Geography and Environment.

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1. Introduction

There is a surging interest in the research of spatial dependence which has already covered a wide range of topics in social sciences;[1] however, so far the focus of empirical work clearly has been on monadic data, while applications in dyadic data are still very rare. The terms "Spatial dependence" or "spatial contagion" refer to a phenomenon in which the behaviour of an actor is influenced by the action of other actors. In dyadic data, the unit of observation is usually a country-pair and the dependent variable describes some form of relationship between these two countries, such as a bilateral flow or a contract. Examples of existing application encompass trade flows (Porojan 2001), preferential trade agreements (Manger 2006), bilateral investment treaties (Elkins et al. 2006, Neumayer and Plümper 2010), double taxation treaties (Neumayer and Plümper 2012), and conflict (Gartzke and Gleditsch 2006). Even though the idea of spatial dependence is not new and there are earlier applications in dyadic data, Neumayer and Plümper (2010) were the first to categorise and name all potential forms of spatial dependence in these two-country-settings. There are different ways to model spatial dependence, namely the spatial-lag, the spatial-x and the spatial error model. In the first model, the weighted values of the dependent variable in all other dyads are included as an independent variable, while in a spatial-x model the dependent variable is regressed on the weighted values of one or more explanatory variables in other dyads and spatial dependence is modeled as a part of the error term in the spatial error model (Anselin 1988). In this thesis, the spatial-lag model is used to test the hypothesis derived from theoretical discussions. In such a model, the linkage between two dyads (or individual members of two different dyads) has to be modelled. Based on Tobler's (1970: 236) observation that "everything is related to everything else, but near things are more related than distant things", which commonly has become known as the first "law" of geography, the most popular link between two units is based on geographical distance, where lower weight is allocated to more remote units. However, the concept of distance can also be transferred to other measures of similarity or channels through which spatial contagion works.

This thesis contributes to the still very limited literature on spatial dependence in dyadic data by analysing three different applications: The second chapter looks at spatial dependence in the diffusion of double taxation treaties, i.e. whether the propensity that a given country pair enters such a treaty depends on the existing tax treaty network. In the

[1] See Plümper and Neumayer (2010) for a list of existing studies.

third chapter, spatial dependence in aid allocation is analysed. Specifically, the question whether the official development assistance of a donor to a given recipient is influenced by aid flows from other donors to the same beneficiary country is addressed. Finally, the fourth chapter on the one hand investigates whether the number of asylum seekers from one source in a destination country depends on the asylum flow from other source countries in the same target country. On the other hand, it is analysed whether the number of asylum applications from one country in a destination country is affected by the number of applications from the same source country in other target countries. In all three case studies, there is comprehensive evidence for spatial dependence which indicates that an implicit assumption of empirical studies which neglect spatial dependence, namely that governments' and individuals' decisions are taken independently of the actions of others, is not appropriate in many cases. In the first case study, spatial contagion provides an explanation for the seemingly irrational behaviour of developing countries when they enter tax treaties with major capital exporters. Bilateral aid and asylum flows have been a popular topic in empirical research. While the potential spatial contagion in asylum flows has been fully neglected so far, existing work on aid neither has modelled the spatial dependence properly nor does it provide theoretical arguments for the specification of the linkages.

This introductory part is intentionally kept short and it outlines only the key concepts and methodological issues since each chapter has its own introduction to ease the flow of reading. In these introductions, the specific topic is set up and the research gap is presented in detail. The remainder of this first chapter is structured as follows: In Section 2, the theory of spatial dependence and its main causes are discussed followed by an illustration of the possible forms of spatial dependence in dyadic data. Section 4 is dedicated to the empirical problems when estimating spatial dependence. Finally, Section 5 provides a brief summary of the theoretical considerations and the main findings of the three applications which form the main part of this thesis.

1.2. Theory of spatial interdependence

There is a plethora of reasons why the action of one individual is influenced by the activities of his or her neighbours, peers, friends or competitors. Nevertheless, Franzese and Hays (2010) describe a few main sources of spatial dependence, which can be used to explain a wide range of observable behaviour. In general, interdependence occurs if the marginal utility of an actor *i* is influenced by the action of another actor *j* (Franzese and Hays 2010). Spatial dependence can be positive or negative. Positive spatial dependence means that contagion leads to a change in the behaviour of the influenced actor in the same direction as the behaviour of the actor that causes contagion. This would be the case, if for instance, one country lowered its corporate tax rate and other countries follow suit. In contrast, negative spatial dependence triggers behaviour of the influenced actor which in the opposite direction of the action of the source of contagion. For example, if one country cuts its greenhouse gas emissions to combat global warming, other countries might reduce their efforts to lower their emissions.

The first main source of spatial dependence is strategic interaction, which is caused by positive or negative externalities. Basically, the externality could either be positive or negative. In the first case, the behaviour of another actor *j* increases the utility of actor *i*. Positive externalities could for example lead to free-riding behaviour, where actor *i* reduces its contribution to a common good if actor *j* provides more of this good. Another example is late mover advantage: this is the case if a company faces lower costs and less risk to enter a market when it is not the first to open up this market. An example for a negative externality is a beggar-thy-neighbour policy, in which the policy action of one country (e.g. a tighter asylum policy) lowers the utility of other countries (e.g. rise in asylum applications). As a consequence, a race-to-the-bottom could evolve in which the influenced actors have an incentive to copy or even magnify the policy action of another country. In general, positive externalities lead to negative and negative externalities to positive spatial dependence (Franzese and Hays 2010). The second basic source of spatial dependence stems from learning from other actors. Learning can also cause both positive and negative spatial dependence. In the first case, actor *i* copies the action of actor *j* after observing that this action proved to be beneficial. Learning could also lead to negative spatial dependence, if actor *i* observes a detrimental effect of an action taken by actor *j* and decides to do the opposite. Finally, also coercion could lead to spatial dependence, most likely in the positive form if actor *j* forces actor *i* to align its behaviour or policy with its own behaviour or policy. On the one hand, such coercion could for example be due to social group pressure among individuals. Governments, on the other hand, may exert coercion on other governments if they have an economic, ideological, social or cultural incentive to do so (Neumayer and Plümper 2012).

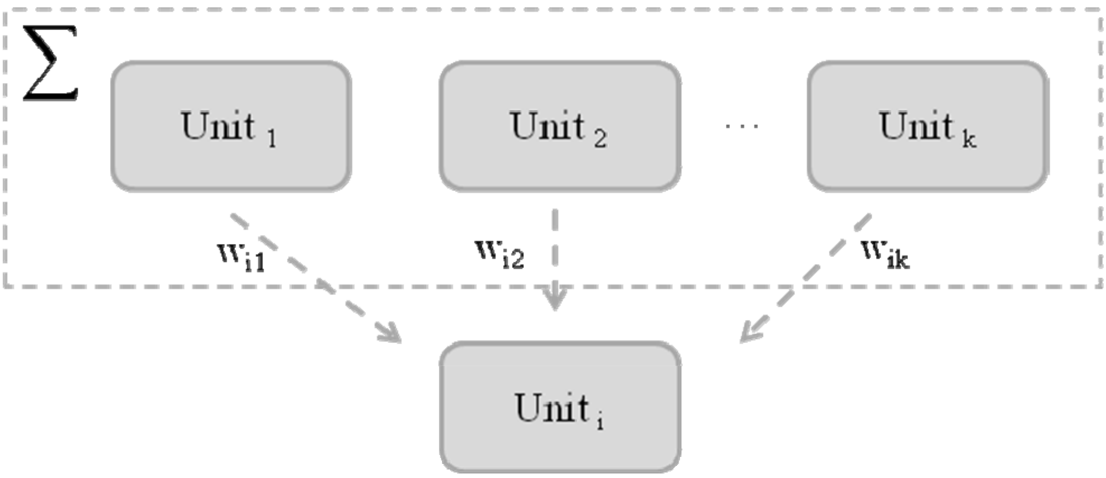
1.3. Modelling spatial dependence

While spatial dependence in monadic data is straightforward and the leeway is limited to the choice of the appropriate weighting matrix, modelling such dependencies in dyadic frameworks allows for a much larger variety of possibilities. After a short review of spatial dependence in monadic data, this section first presents all possible forms of spatial contagion in dyadic frameworks and then discusses the options for the weighting matrix. Much of this work can be credited to Neumayer and Plümper (2010) who were the first to address spatial dependence in dyadic data.

Forms of spatial dependence

If spatial contagion exists, in monadic data the outcome in the observational unit *i* depends on the weighted actions of other units *k*. In social sciences such an outcome could be the policy choice of a local or national government, the advertising budget of a company or the decision of an individual to buy a certain product, which is influenced by the policy choices of other governments, the money spent by other firms or the buying behaviour of friends. The linkage between the unit under observation and the other units is modelled in a weighting matrix w_{ik}, which connects unit *i* to other units *k*. This relationship is depicted in Figure 1.

Figure 1: Spatial dependence in monadic data



Source: Own illustration.

Abstracting from all control variables and a potential time-dimension, the spatial lag model for monadic data can be formally expressed as follows (Neumayer and Plümper 2010):

$$y_i = \rho \sum_k w_{ik} y_k + \varepsilon_i \tag{1}$$

where the dependent variable y_i is the outcome in unit *i* which is explained by the weighted sum of the outcomes in other units *k* and ρ is the spatial lag parameter to be estimated. Y_i is said to be spatially dependent on y_k if the spatial lag parameter is statistically significant different from zero. In general, a positive spatial lag parameter indicates a complementary relationship between the unit under observation and other units, whereas a negative parameter shows that they are substitutes.

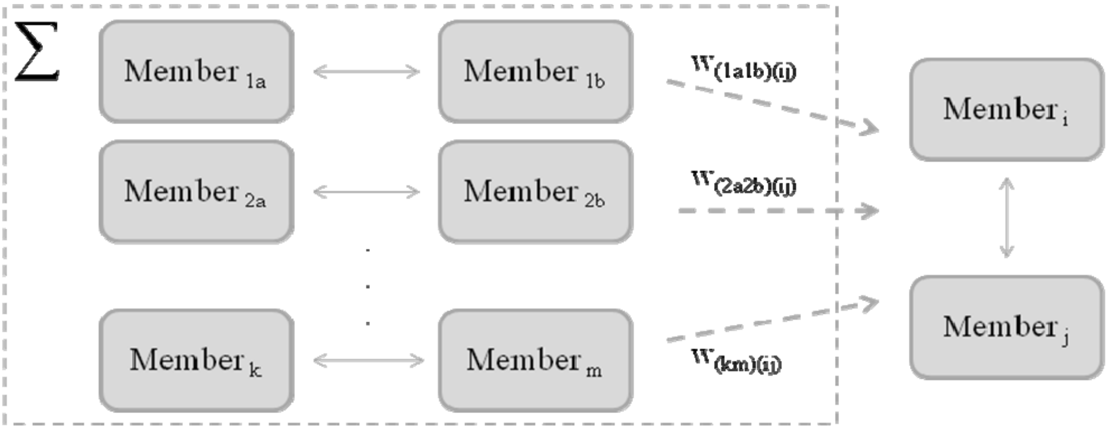
In dyadic data generally two different forms can be distinguished, namely undirected and directed dyads. In directed dyads, there is a clear distinction between the sender and the recipient, the exporter and the importer, or the aggressor and the victim, whereas such a differentiation is not possible or not theoretically interesting in undirected dyads.[2] Examples for directed dyads are international flows such as trade, foreign direct investment or migration, but also military actions and telecommunication connections, while a voluntarily entered international agreement is an example of an undirected dyad. Modelling spatial dependence in undirected dyads closely resembles the spatial lag-model in monadic data with the notable difference that not units depend on each other but dyads. This relationship is shown in Figure 2. Member *i* and member *j* are the two countries forming the dyad under observation and the double-headed arrow indicates that no direction of their interaction can be distinguished. The relationship in dyad *ij* (or dyad *ji*) spatially depends on the weighted aggregate relationship between other dyads *km* (or *mk*), where the weight $w_{(km)(ij)}$ determines the influence of a given dyad and represents the connectivity between dyad *km* and dyad *ij*. The weights are illustrated as dashed arrows in the figure. Formally, this reads as follows (Neumayer and Plümper 2010):

$$y_{ij} = \rho \sum_{km \neq ij} w_{pq} y_{km} + \varepsilon_{ij} \tag{2}$$

[2] One can still distinguish dyad member *i* from dyad member *j*, but dyad *ij* is the same as dyad *ji*.

where w_{pq} is a more general form of the weighting matrix connecting dyad *ij* and dyad *km*.[3]

Figure 2: Spatial dependence in undirected dyads

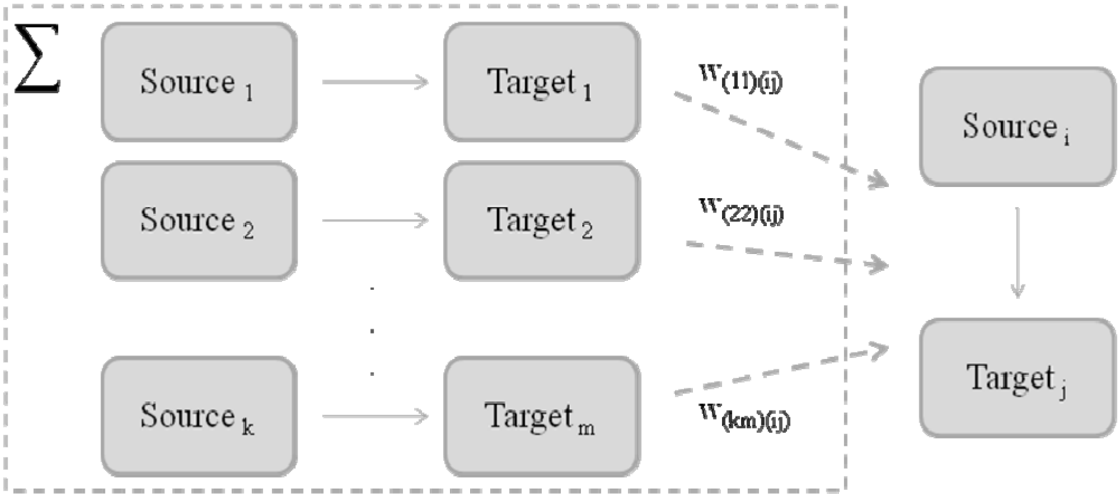


Source: Own illustration.

In directed dyads, there are five different forms of spatial dependence. The one that is equivalent to undirected dyad contagion is named directed dyad contagion by Neumayer and Plümper (2010) and illustrated in Figure 3. Here, for example, the trade flow between source *i* and target *j* depends on the weighted trade between other dyads *km* and the weighting matrix $w_{(km)(ij)}$ again represents the connectivity between the influencing dyads and the dyad under observation. The formal expression of directed dyad contagion is the same as in undirected dyad contagion, since in both cases a dyad depends on a dyad and the only difference is the directedness.

[3] Note that for clearness, Figure 2 shows a situation in which dyads containing either member *i* or member *j* have no contagious effect, whereas the mathematical expression only excludes the dyad *ij*, but for instance not the dyads *ki* and *mj*.

Figure 3: Spatial dependence in directed dyads: directed dyad contagion



Source: Own illustration.

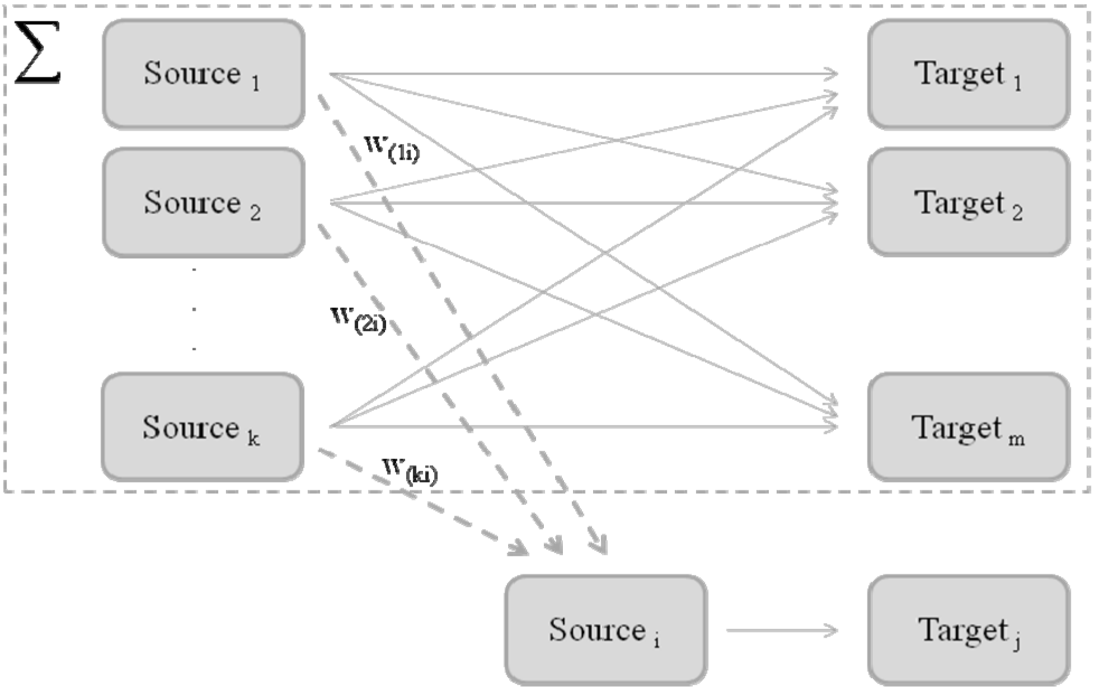
The four remaining forms of spatial dependence are slightly more complex. The basic modelling decision is whether the aggregate action of other sources or targets matters or their interaction with a specific target or a specific source. Staying with the example of international trade, the first type is aggregate source contagion, where the trade volume between exporter *i* and importer *j* depends on the weighted sum of the trade between all other exporters *k* and all other importers *m*. In Figure 4, trade flows are depicted as solid arrows.[4] Under aggregate source contagion, the weighting matrix represents the connectivity between sources, more specifically between other sources *k* and the source under observation *i*. The weights are again pictured as dashed arrows in the graph.

Following Neumayer and Plümper (2010), the most parsimonious estimation model for aggregate source contagion reads as follows:

$$y_{ij} = \rho \sum_{k \neq i} \sum_{m} w_{pq} y_{km} + \varepsilon_{ij} \tag{3}$$

[4] Again, for reasons of clarity trade involving either source *i* or target *j* is excluded in the graph, but this is not necessarily warranted, as such a choice must be justified theoretically.

Figure 4: Spatial dependence in directed dyads: aggregate source contagion

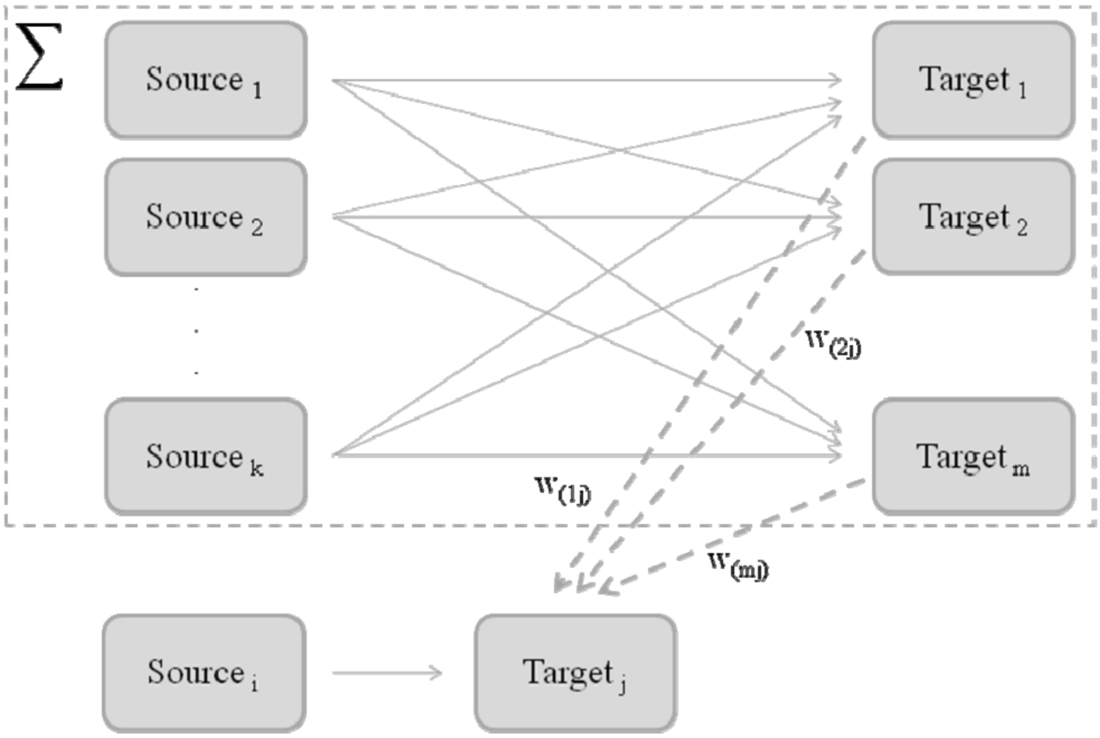


Source: Own illustration.

The counterpart to aggregate source contagion is aggregate target contagion (Figure 5), where the trade flow between source *i* and target *j* again depends on the weighted sum of the trade flows between all other sources *k* and all other targets *m*. However, here the weighting matrix measures the connectivity between target *k* and target *j*. The formal expression of aggregate target contagion is (Neumayer and Plümper 2010):

$$y_{ij} = \rho \sum_{k} \sum_{m \neq j} w_{pq} y_{km} + \varepsilon_{ij} \tag{4}$$

Figure 5: Spatial dependence in directed dyads: aggregate target contagion

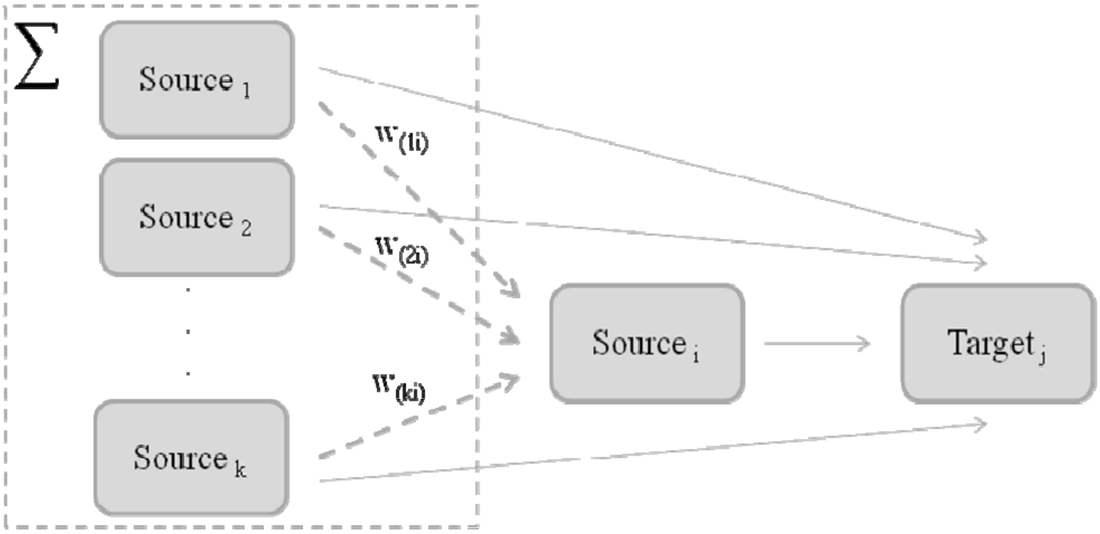


Source: Own illustration.

Finally, in the last two forms of spatial dependence only a subset of sources or targets influences dyad *ij*, namely those sources and targets which are linked to target *j* and source *i*, respectively. First, under specific source contagion (Figure 6), the trade volume between source *i* and target *j* depends on the weighted sum of trade between all other sources *k* and the same target *j*. The weighting matrix represents the connectivity between source *k* and source *i* and the formal expression is as follows (Neumayer and Plümper 2010):

$$y_{ij} = \rho \sum_{k \neq i} w_{pq} y_{kj} + \varepsilon_{ij} \tag{5}$$

Figure 6: Spatial dependence in directed dyads: specific source contagion

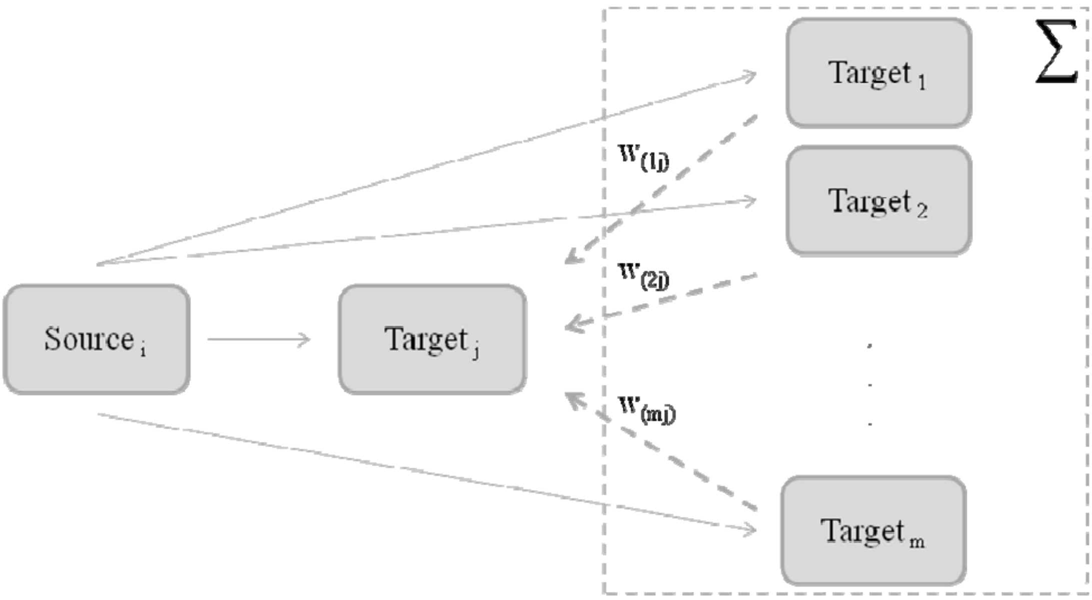


Source: Own illustration.

The second specific form is specific target contagion, under which the trade flow between exporter *i* and importer *j* is influenced by the weighted aggregate of trade between source *i* and other targets *m* (Figure 7). The weighting matrix measures the connectivity between other targets *m* and the target under observation *j*. The mathematical expression for specific target contagion is (Neumayer and Plümper 2010):

$$y_{ij} = \rho \sum_{m \neq j} w_{pq} y_{im} + \varepsilon_{ij} \quad (6)$$

Figure 7: Spatial dependence in directed dyads: specific target contagion



Source: Own illustration.

The choice of the weighting matrix

In general, the weighting matrix is used to model the connectivity between units in monadic data, between dyads under undirected and directed dyad contagion and between sources or targets in the remaining four forms. This section addresses only the key issues, a more comprehensive discussion can be found in Neumayer and Plümper (2010) and on the row-standardisation in Plümper and Neumayer (2010) as well as in Neumayer and Plümper (2012). Based on Tobler's (1970) first law of geography, the most common weighting matrix is based on geographical distance (Beck et al. 2006), in which the allocated weight decreases with distance or the influence is limited to adjacent units. However, spatial influence often is not only determined by geographical distance, but it could also be due to a common membership in an international organisation or by the degree of structural similarity of countries. Such dependencies can also be modelled in the weighting matrix. Independent of whether the data is monadic or dyadic and the directedness of the dyad, the connectivity in the weighting matrix itself can be directed or undirected (Neumayer and Plümper 2010). Geographical distance, export product similarity and common membership are examples of an undirected connectivity. In this case the weighting matrix w_{ik} is identical to w_{ki} due to the undirectedness of the connectivity. An example for a directed connectivity is imports, if a country is influenced by other countries and the extent of this influence is determined by the import volume from a given exporter. Here, with exports as the connectivity variable, the weighting matrix w_{ik}, in which case connectivity measures exports from country i (the country under observation) to country k, is different from the weighting matrix w_{ki}, where connectivity measures exports from country k to country i. Such a simple form of the weighting matrix is appropriate if the connectivity between single countries is modelled rather than between dyads, for instance in monadic data or between sources in specific and aggregate source contagion. If the connectivity in dyadic data is not modelled between the member of one dyad and the member of another dyad, but between dyads as a whole, the weighting matrix gets more complicated. Basically, there are two main options. First, the linkage between one dyad and another is still represented by the connectivity between individual members of different dyads, but these linkages have to be combined to form the connectivity between two dyads. For instance, if theory suggests that all countries with a common border exert the same amount of influence, the weighting matrix between dyad member i of the dyad under observation and the dyad member k of other dyads contains the value of one if i and k share a common border. Similarly, the corresponding cell of the weighting matrix between dyad member j of the dyad under observation and dyad member m of other dyads gets the value of one, if j and m are adjacent countries. These two weighting matrices have to be combined into one to represent the connectivity between dyad *ij* and dyad *km*. The two most straightforward ways of doing this are either summing the two matrices up or multiplying them. If the weighting matrices are added up, a substitutive relationship between the two linkages is assumed. In this case, the weighting matrix $w_{(ik+jm)}$ contains a one if either country *i* shares a border with country *k* or country *j* has a common border with country *m*.[5] In contrast, both linkages are regarded as complements, if the product of both weighting matrices is taken. Here, the weighting matrix $w_{(ik*jm)}$ contains a one only if both country *i* shares a border with country *k* and country *j* is adjacent to country *m*. Second, and alternatively, the weighting matrix could consist of a link between a property of the dyad *ij* itself on the one hand and the dyad *km* on the other, in which case the weighting matrix would be $w_{(ij)(km)}$.

The second important question regarding the weighting matrix is whether to row-standardise or not. In a row-standardised matrix, each cell of the matrix is divided by the sum of its row. As a consequence, the spatial lag is no longer the weighted sum of the lagged dependent variable in other units or dyads but the weighted average. Row-standardisation has several advantages for the interpretation of the results: First, the spatial lag has the same metric unit than the dependent variable itself, which allows a direct comparison of the coefficient of the spatial lag with the coefficient of a temporal lag (Ward and Gleditsch 2008). Second, it enables the researcher to interpret the coefficient size of the spatial lag as the approximate strength of interdependence (Franzese and Hays 2008). And third, the stationarity requirement can be checked easily if the sum of the coefficient of the spatial lag and the coefficient of the temporal lag is less than one (Franzese and Hays 2008). While row-standardisation is common practice in spatial econometrics (Anselin 2002, Francese and Hays 2006), Plümper and Neumayer (2010) argue that this procedure has to be theoretically justified as it changes the relative weight that is given to other observations, unless the special case of a unitary weighting matrix is used, which gives equal weight to all other observations. Staying with the example of a contiguity weighting matrix, Portugal has only one neighbour, namely Spain, whereas the Netherlands have two neighbours (Belgium and Germany). Without row-standardisation, Spain exerts the same influence on Portugal as Belgium and Germany on the Netherlands. In a row-standardised weighting matrix, however, Belgium and Germany both exert 50 percent of the total weight

---

[5] In this case of a binary weighting matrix, the combined matrix contains the value of two if both country *i* and country *k* and country *j* and country *m* share a border. As a consequence, the influence of such a dyad would be twice as high as the impact of a dyad in which only one dyad member has a common border. The researcher has to decide whether this is theoretically justified.

on the Netherlands, while Spain's influence is a 100 percent of all impact on Portugal. This is because row-standardisation removes all level effects in the weighting matrix (Neumayer and Plümper 2012).

1.4. Empirical problems to identify spatial dependence

Even if spatial patterns are found in a regression analysis, i.e. the coefficient of the spatial lag is statistically different from zero, these effects do not necessarily represent spatial contagion, but they could also be due to spurious effects. The reason is that geographically close units are likely to be more similar than more remote units, as many observable and unobservable phenomena such us cultural habits, preferences or institutions are spatially clustered (Plümper and Neumayer 2010). If these influencing factors can be observed, they are referred to as spatial clustering. They are less problematic in an empirical analysis if appropriate measures or proxies allow controlling for them. More challenging are unobserved factors which are spatially correlated and are known as unobserved spatial heterogeneity. If these factors are correlated both with the outcome and the spatially lagged dependent variable, the latter is correlated with the error term and hence endogenous as one of the main assumptions of the OLS model is violated. In this case, the estimated coefficients are biased. Distinguishing spatial clustering and unobserved spatial heterogeneity from causal effects is widely known as Galton's Problem (Galton 1889).[6]
Apart from observable and unobservable factors that are spatially correlated, also common trends and common shocks could lead to spurious spatial dependence. A common trend exists if there is a factor which varies over time in numerous units that are connected via the weighting matrix. For instance, population grows in nearly all countries, even though it is likely that there is no contagious effect of population growth in one country on the fertility of the population in other countries. However, the inclusion of the weighted population growth rate in neighbouring countries would probably indicate a positive effect. Similarly, a common shock is a one-time event which affects the outcome in more than one unit and thereby both the dependent variable in a given unit and the spatial lag. For instance, a wide-spread drought might affect child mortality in geographically close countries without representing a contagious effect. Both common shocks and common trends lead to upward biased estimates of the spatial lag.

[6] Some authors, such as Gibbons and Overman (2010), argue that distinguishing these effects and hence identification is nearly impossible and advocate exploiting variance caused by "natural experiments". While such an approach would render internal validity more likely, it considerably limits external validity if the results are not generalisable.

Plümper and Neumayer (2012) suggest several steps in model specification to mitigate the influence of common shocks and trends and to alleviate the problem of unobserved spatial heterogeneity, hence, to make spurious spatial effects less likely. First, and most importantly, the inclusion of dyad-level fixed effects helps to mitigate the problem of unobserved spatial heterogeneity as they capture all time-invariant factors that are specific to a dyad. While this has the drawback that no estimates for time-invariant control variables can be obtained, it has the big advantage that all unobservable factors which do not vary over time are automatically controlled for. Estimation of the coefficients is based on over-time variation of the variables only. This means that identification of spatial effects rests on the much weaker assumption that there is no unobserved spatial heterogeneity in *changes*, rather than in *levels*. Second, the model specification should be as broad as possible to account for spatial clustering in observable variables. If these variables are excluded from the regression but are correlated with the spatial lags, the coefficient estimates for the spatially lagged dependent variable suffer from an omitted-variable bias. Third, the inclusion of a temporal lag of the dependent variable captures common trends and also accounts for temporal dynamics. Finally, the inclusion of a *t-1* set of year dummies controls for any general global trend in the dependent variable which leads to the common trend phenomenon described above, but it also captures the effect of common shocks. However, modelling such a general time trend assumes that the year effects are the same for every dyad in the sample which might not adequately address the common trend problem if the trend is specific to a subset of dyads. A more flexible way of modelling is the inclusion of dyad-specific time trends, or even time-trends that are specific to one or both of the dyad members. While these trends effectively control for common trends, they aggravate the problem of multicollinearity especially with trended variables such as population size.

1.5. Summary of the findings

This thesis provides three applications of spatial dependence in dyadic data, namely in the diffusion of double taxation treaties, in the patterns of bilateral and multilateral aid giving and in asylum migration. While these examples differ in some terms, they share certain important common characteristics. Starting with the differences, the first is an example of an undirected dyad, whereas the latter two are directed dyads. In the first two applications, the dependent variable is a government choice, while in the last it is the aggregate of individuals' decisions to migrate. The most prominent common feature is that all three are based on dyadic data, specifically on country-pairs. In all case studies it is analysed, how policy choices or migration flows in one country are affected by the actions of - or the migration patterns in - other countries. More specifically, each example addresses the relationship between developing and developed countries, even though from a different angle of view: the first example - double taxation treaties - analyses how strategic interaction between developing countries can explain their seemingly irrational behaviour. The second chapter examines how the aid allocation decision of a donor to a developing country is influenced by the aid amount that other donors give to the same recipient. Finally, the third case study scrutinises whether the number of asylum seekers from one source country in a destination country depends on the size of the asylum flows from other sources and to other targets. Here, developing countries are the main countries of origin of asylum migration, while the flows are often directed to industrialised nations. In the following, the research gap, the main theoretical arguments and the findings of each chapter are presented.

Spatial dependence in the diffusion of double taxation treaties

Double taxation treaties (DTTs) are international contracts which are concluded between two independent tax jurisdictions with the main task of allocating the taxation rights between the two signatory states. By alleviating twofold taxation, the aspired goal of such treaties is to facilitate the exchange of goods and services and to promote the exchange of capital and persons. Subordinate goals are administrative collaboration of tax authorities, facilitating the enforcement of domestic tax laws, combating tax evasion and the enshrinement of fiscal investment incentives which can be neutralised under certain circumstances in the international tax regime. Another advantage, particularly relevant for developing countries, is that tax treaties might be regarded as a commitment to jurisdictional predictability and stability by potential investors. Finally, tax treaties provide a global standard which lowers communication and enforcement costs and leads to positive network externalities if a large number of countries entered such a treaty. Concluding such a treaty, however, can lead to sizeable costs. The first cost bloc is administrative resources which are tied up during the negotiation and ratification process and which can be tremendous particularly if the tax systems of the signatory states differ widely and different languages have to be synchronised. The second and even more important cost factor is the potential loss of tax revenues for one of the contracting states. The vast majority of tax treaties are based on the OECD model treaty. These DTTs strongly favour residence- over source-taxation. This means that in the case of foreign direct investment, more taxation rights are granted to the residence country of the investor than to the source country in which the taxable profit is generated. These source countries (in terms of profit) are left with a withholding tax usually ranging between zero and 15 percent which is specified in the treaty and credited against the tax liability of the investor in its country of residence to avoid double taxation. This does not impose a burden on the domestic tax revenues if the bilateral FDI stocks between the two contracting states are roughly equal as benefits granted in the own country to an investor from the other state should be offset by benefits offered to the own investors in the other state. In practice, however, bilateral FDI stocks are often highly asymmetric, where one country is a net-capital exporter and the other a net-capital importer. In such a constellation, the net-importer faces a considerable net-loss in tax revenues which is determined by the level of domestic tax rates, the withholding tax rate and the degree of asymmetry in the FDI stock. Developing countries are not always but quite often in the position of a net-capital importer if the treaty is concluded between a developing and an industrialised country. Even though these countries might be worse off by signing a DTT, treaties between a developing and a developed nation are a widespread phenomenon. While theoretical arguments have been brought forward to explain this seemingly irrational behaviour, the analysis presented in chapter two is the first to address this issue empirically.

To start with, there are two arguments for developing countries to enter a DTT which do not lead to spatial dependence in the diffusion of tax treaties: First, fiscal incentives in the form of tax reductions or tax holidays which are provided by a country to attract FDI can be neutralised if the country of residence of the investor grants tax credits for the amount of taxes paid in a foreign country to unilaterally avoid double taxation. In this case, a lower tax rate in the host country has no effect as this leads to less tax credits in the home country and does not affect the overall tax burden of a company. To avoid this, many DTTs contain a tax-sparing provision under which the original amount of tax is credited in the home country. The second argument for net-capital importing countries to enter a tax treaty is the dynamic inconsistency problem. The tax jurisdictions particularly of developing countries might be perceived as unreliable by investors. In fact, due to the fact that FDI is at least partly irreversible, the host country has an incentive to attract investors by lower tax rates and raise taxes after the investment took place. Since provisions for the maximum host country tax rate are included in a DTT, such a treaty mitigates the dynamic inconsistency problem. The final two arguments lead to spatial dependence, one through positive externalities and the other through strategic interaction among host countries. A dense network of tax treaties that are based on the same model treaty lowers communication and enforcement costs not only for existing treaties, but also for newly entered DTTs. This positive network effects make the conclusion of an additional tax treaty more likely, the larger the network of existing treaties already is. The most important argument to explain the conclusion of a tax treaty in an asymmetric dyad, however, is strategic interaction among host countries. Baistrocchi (2008) describes the situation net-capital-importers find themselves in as a prisoner's dilemma: Host countries are usually keen to attract FDI and tax treaties are perceived as one way to make the own country a more attractive investment location. If a country signs a DTT with a major capital exporter, it gains a relative advantage over its competitors and these have an incentive also to enter a tax treaty with the same capital exporter to offset their relative disadvantage. As a consequence, no host country has an advantage (in terms of FDI attractiveness) of entering a tax treaty, but net-capital-importing countries would be better off, if none signed a DTT with a major capital-exporter. By deteriorating the relative competitive position of other countries, a net-capital-importer which enters a DTT imposes a negative externality on other countries.

To test the argument that DTT diffusion is driven by the behaviour of competitors for FDI, the treaty formation between 186 countries in the period from 1969 to 2005 is analysed. The sample covers more than 2,300 treaty conclusions. Since there are theoretical arguments for both the capital-exporter (maximise tax revenues) and the capital-importer (attract FDI) to start the treaty-negotiations and such information is not available for many treaties, DTT conclusion is modelled as an undirected dyad. To delineate competing countries, three different concepts are used: First, all countries in the same region are assumed to compete for the same FDI. Second, countries which export a similar basket of goods and third, countries that serve similar export markets are taken as competitors. The main estimation technique is a Cox proportional hazard model, which provides estimates for factors influencing the time until a tax treaty is signed in a dyad.

In line with the theoretical expectations, robust evidence is found that the treaty conclusion between two countries is positively influenced by tax treaties that are negotiated by competing countries. However, while competitive pressure is exerted by countries that export similar goods and from countries in the same region, no effect can be established for countries which serve similar export markets. This indicates that countries with a similar production structure are regarded as competitors for foreign capital as countries strive to attract a certain type of capital rather than a broad range of investors which export finished products to specific markets. It is also found that the degree of spatial dependence systematically varies across countries: Fist, the spatial effect decreases with the number of DTTs a country already has signed. If a country already has a dense DTT network, likely it has covered all major capital exporters and is less influenced by the actions of other countries. Once a country gains independence, it is able to enter its own tax treaties. It is argued that such a newly independent country faces a higher competitive pressure as it has to neutralise its relative disadvantage. In line with this argument, spatial dependence is found to decrease with the years from independence.

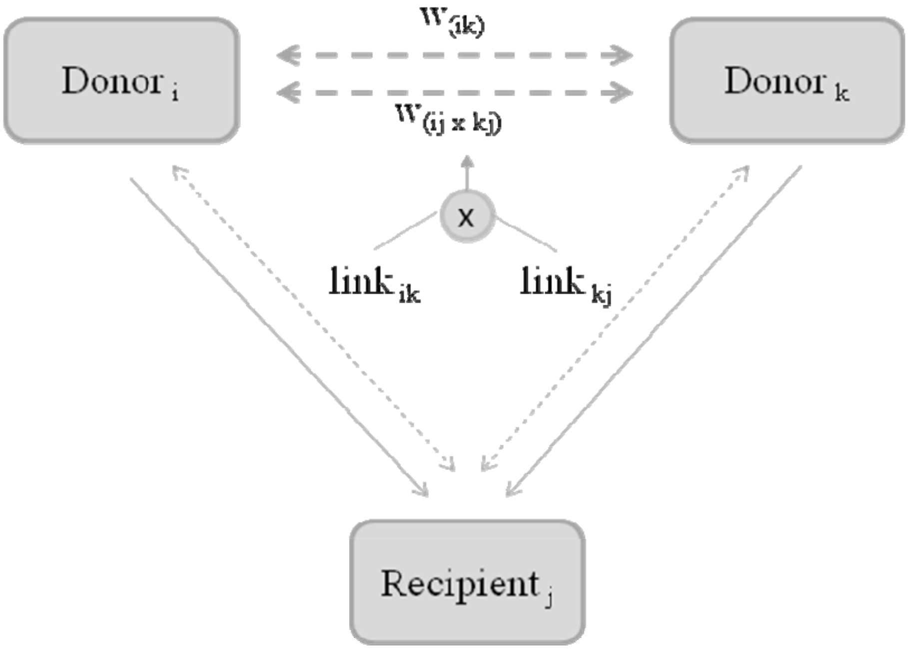
The consequences of such behaviour can be far-reaching for developing countries, which find themselves frequently in the position of a net-capital-importer: First, they lose tax revenues. Second, the effectiveness of a DTT to gain a competitive edge in the rivalry for foreign capital is lost once other competitor countries also have signed DTTs with major capital exporters. Collective action among developing countries could solve this dilemma, but is unlikely given the large number of parties involved and due to the fact that no single country could gain an advantage with a multilateral treaty. Also model treaties which are more favourable for developing countries failed to gain influence in international taxation.

Spatial dependence in bi- and multilateral patterns of aid giving

The third chapter deals with the question, whether the decision by a donor how to allocate its official development assistance (ODA) is influenced by the aid allocation decisions of other bilateral and multilateral donors. There are theoretical reasons for both positive and negative spatial dependence. In the first case, a given donor will increase (decrease) its aid to a given recipient if other donors give more (less) aid to the same recipient. Negative spatial dependence is found if a donor reduces (raises) the aid budget to a recipient country if other donors allocate more (less) aid to this recipient. The first reason for positive spatial dependence is the uncertainty necessarily involved in aid projects, e.g. whether the money actually reaches the beneficiary, how the money will be spent and whether the envisaged development effects will materialise. Since aid authorities are accountable to the national taxpayer, donors might follow the example of other donors and give funds to the same recipient countries since their allocation might be a signal for good aid projects (Vázquez 2008). The second cause for positive spatial dependence is strategic interaction among donors. If aid is at least partly used to pursue national economic, political and military interests of the donor, each donor has to closely observe the actions of other donors and react correspondingly. For instance, if another donor allocates a larger budget to an important trade partner and this is also an important trade partner of the donor under observation, this donor should follow suit and increase its aid budget in the recipient country to avoid losing influence. Negative spatial dependence could be a sign of donor specialisation over time if each donor concentrates its activities on a limited set of countries and this set is not identical for all donors. Such a specialisation would be in line with various political pledges to improve donor coordination. Taken together, either learning from others or strategic interaction form the cause of spatial dependence in this context.

While the analysis provided in Chapter 3 is not the first to include the aid allocation by other donors into an estimation model, it is the first to model the spatial dependence explicitly by using appropriate weighting matrices. All former studies simply add the average or aggregate aid by other donors as a control variable, without providing theoretical arguments for this specification choice. They also never talk about "spatial dependence" which corroborates the assumption that the authors are not consciously analysing spatial dependence. The empirical work in Chapter 3 makes use of a newly published aid dataset and covers 20 bilateral donors and three multilateral donor organisations, 139 recipient countries and spans the period from 1974 to 2008. The aid allocation decision is modelled as a two-step process and spatial dependence is analysed in both steps: in the first step, a donor decides to which recipient countries a positive amount of aid is allocated in a given year (eligibility stage) and in the second step, the actual aid budget is distributed to the countries on the recipient list (allocation stage). Aid allocation is a classical example of a directed dyad. The type of spatial dependence analysed is specific source contagion, however, some of the weighting matrix used to model the connectivity between different sources (donors) take their relationship with the target (recipient) into account. This is illustrated in Figure 8, in which for reasons of simplicity only two donors are shown. As in Section 3, the solid lines show the flow from the source to the target country, in this case the aid from donor *i* and donor *k* to recipient *j*. The first set of weighting matrices $w_{(ik)}$ is straightforward and models the direct connectivity between donor *i* and donor *k*. The empirical specification includes weighting matrices which allocate equal weight to all other donors, equal weight to members of a specific subgroup of donors and weight according to the importance of a given donor, measured by its share at the global aid budget. In contrast, the weights in the second set $w_{(ij\ x\ kj)}$ depend on the bilateral relationship both between donor *i* and recipient *j* and between donor *k* and recipient *j*. The product of these two relationships is taken to form the weight between donor *i* and donor *k*. The relationships which are tested in the empirical work encompass trade relations, diplomatic relations and military strategic relations.

Figure 8: Specific source contagion in Chapter 3



Source: Own illustration.

In the first stage, the eligibility stage, comprehensive evidence for positive spatial dependence is found. A donor is more likely to give aid to a specific recipient, if other donors also allocate some positive amount of aid to this recipient. The same effect holds true if the recipient-specific weighting matrices are used. Also in the second stage, all spatial lags are positive and statistically significant as long as they are estimated individually. If they are all included in one estimation model, only the spatial lag which allocates weight to donors according to their relative importance and the spatial lag accounting for military strategic relationships remain statistically significant. If different donor groups are analysed separately, it can be seen that the above results are representative for bilateral donors in general and the big Western donors in particular, whereas no evidence is found for spatial dependence among multilateral institutions. Interestingly, for the group of like-minded donors, Canada, Denmark, the Netherlands, Norway and Sweden, which have the reputation that the promotion of good governance and democracy plays an important role in their aid allocation decisions, a different result is found: these donors generally allocate more aid to recipients which receive more aid from all other donors and particularly from other like-minded donors.

The results suggest that there is a considerable degree of spatial dependence in aid allocation – but only very limited evidence for strategic interaction among donors; at least not to pursue their economic and diplomatic goals. In their decision to allocate aid, donors clearly follow the aid allocations of other donors and they take aid allocation decisions of large donors as a guideline for the distribution of their own aid budget.

Spatial dependence in asylum migration

Finally, the fourth chapter deals with spatial dependence in international asylum migration flows. Two different forms of spatial dependence are analysed, namely specific source contagion and specific target contagion. The first type means that the number of asylum applications from one source country in a target country depends on the number of applications from other source countries in the same destination country. In contrast, under specific target contagion is the number of asylum seekers from a source country in a target country influenced by the number of applications from the same country of origin in other destination countries. Migration in general and asylum migration in particular is subject to a great deal of uncertainty, as the migrant has to leave the familiar surroundings and social ties and accommodate to a new environment. Personal networks reduce the risk of migration since the members can assist in the search for accommodation or employment, help with language problems or provide other benefits such as a feeling of belonging. The importance of such networks is already well documented in the empirical and qualitative literature; however, so far it has always been assumed that these personal networks are confined to nationals of the migrant. This is not necessarily the case if also migrants from other adjacent or geographically close source countries facilitate transition and thereby reduce migration costs. If this is true, a higher number of asylum migrants from other countries in a given destination country should lead to more asylum seekers from the source country in question in the same target country. The second argument for specific source contagion is that there are scale effects in the international people smuggling networks. Many destination countries have strict border controls to combat illegal immigration and to deter asylum seekers. This is why the large majority of asylum migrants rely on the services of human traffickers. These agents invest time and money to make contacts with carriers, find suitable routes and loopholes to cross the border of the target country. If main routes have been established between a major source and a destination country or region, getting connected to the network is cheaper for surrounding countries via feeder routes. This should lower transport costs for migrants from geographically close countries and make their numbers dependent on the size of the asylum flow from a major source country. Learning from role-models and the positive externality provided by other migrants are the main cause of specific source contagion.

Policymakers in destination countries use a wide range of measures to discourage asylum seekers from lodging their application in their country, such as border controls, visa restrictions and sanctions against carriers which transport people without a valid visa, penalties for human traffickers, lower recognition rates, and restrictions on employment or limited welfare benefits. This is not only because of the fiscal costs associated with hosting asylum seekers, but also because of strong public resentments against asylum migrants, even though these are generally more directed against bogus refugees. These measures increase migration costs and could thereby deter some people from leaving their home country. However, most will either try their luck and lodge an application despite of a restrictive asylum policy or immigrate without applying for asylum and stay underground. Yet, some asylum seekers might also decide to file their application in a more lenient target country, i.e. they are not deterred, but only deflected. In this case, a lower number of asylum seekers from a source country in one target country should be associated with higher number of asylum migrants from the same source country in other targets, which means that there is negative specific target contagion. Here, deflecting asylum seekers to other countries represents a negative externality as the main source of specific target contagion as a form of spatial dependence.

Both forms of spatial contagion are tested using a dataset which includes the bilateral number of asylum applications from 138 sources in 21 destination countries over the period 1982 to 2008. To analyse specific source contagion, several weighting matrices based on geographically proximity of the source countries are used. Furthermore, to test whether personal networks are also based on a common language, an appropriate weighting matrix is employed. For specific target contagion, it is tested whether asylum seekers are attracted from more restrictive destination countries and deflected to countries with a more liberal asylum policy.

The estimation results show comprehensive evidence for positive specific source contagion. This is in line with both theoretical explanations outlined above. The number of asylum seekers from a given source in a target country is positively influenced by the number of asylum seekers from other source countries which share a common border with the source country in question, source countries from the same region or source countries in which the same language is spoken. The same holds true if all other source countries are weighted according to their geographical proximity to the country of origin under observation. If all spatial lags are estimated together, however, the spatial lag with the common language weighting matrix loses significance. This indicates that there is no personal network effect from migrants from other source countries with the same language, which are not geographically close to the source country in question. For specific target contagion, the expected negative spatial effects are found, i.e. asylum seekers are in fact deflected to countries with a more lenient policy and attracted from countries with a tighter asylum regime.

Both findings have some important policy implications: It has been argued that personal networks originating from migrants from the same source country limit the effectiveness of deterrence policies, as a larger stock of migrants already in the country reduces the costs for each additional migrant (Hatton 2004). As a consequence, if a host country starts tightening asylum policy as a consequence of a mass influx, this step might be too late as those already in the country are not affected by the policy changes. A similar argument applies to migrants from other source countries if deterrence measures are targeted against migrants from a specific source country. Migrants from other source countries are not affected, but provide a benefit to asylum seekers from a given source country and lower thereby their migration costs. Hence, the effectiveness of targeted policy measures is curtailed. On the other hand, a tighter asylum policy provides a negative externality to other destination countries as these face higher application numbers as a consequence. They in turn have an incentive to react with tightening their policy which fuels a race-to-the-bottom in asylum standards. This negative externality provides one more reason for policy harmonisation and coordinated approaches to share the burden – not only, but particularly in order to guarantee a minimum treatment to refugees.

References

Anselin, Luc (1988), *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer.

Anselin, Luc (2002), Under the hood: Issues in the specification and interpretation of spatial regression models, *Agricultural Economics*, 27(3), pp. 247-267.

Baistrocchi, Eduardo (2008), The Use and Interpretation of Tax Treaties in the Emerging World: Theory and Implications, *British Tax Review*, 4, 352-391.

Beck, Nathaniel, Kristian Skede Gleditsch, and Kyle Beardsley (2006), Space is more than geography: Using spatial econometrics in the study of political economy, *International Studies Quarterly*, 50, pp. 27-44.

Elkins, Zachary, Andrew Guzman and Beth Simmons (2006), Competing for Capital: The Diffusion of Bilateral Investment Treaties, 1960-2000, International Organization, 60(4), pp. 811-846.

Franzese, Richard and Judith Hays (2006), Strategic interaction among EU governments in active-labor-market policymaking: Subsidiarity and policy coordination under the European Employment Strategy, *European Union Politics*, 7(2), pp. 167-247.

Franzese, Richard and Judith Hays (2008), *Spatial-Econometric Models of Interdependence*, book prospectus.

Galton, Francis (1889), Discussion on Edward B. Taylor's "On a Method of Investigating the Development of Institutions, Applied to Laws of Marriage and Descent", *Journal of the Anthropological Institute of Great Britain and Ireland*, 18, 270-272.

Gartzke, Erik and Kristian Gleditsch (2006), Identity and Conflict: Ties that Bind and Differences that Divide, *European Journal of International Relations*, 12(1), pp. 53-87.

Gibbons, Steve and Henry Overman (2010), Mostly Pointless Econometrics?, *SERC Discussion Paper 61*, Spatial Economics Research Centre: London.

Hatton, Timothy (2004), Seeking Asylum in Europe, *Economic Policy*, 19(38), pp. 5-62.

Manger, Mark (2006), *The Political Economy of Discrimination: Modelling the Spread of Preferential Trade Agreements*, Working Paper, Montréal, Canada: McGill University.

Neumayer, Eric and Thomas Plümper (2010), Spatial Effects in Dyadic Data, *International Organization*, 64(1), pp. 145-166.

Neumayer, Eric and Thomas Plümper (2012), Conditional Spatial Policy Dependence: Theory and Model Specification, *Comparative Political Studies*, forthcoming.

Plümper, Thomas and Eric Neumayer (2010), Model Specification in the Analysis of Spatial Dependence, *European Journal of Political Research*, 49(3), pp. 418-442.

Porojan, A. (2001), Trade Flows and Spatial Effects: The Gravity Model Revisited, *Open Economics Review*, 12(3), pp. 265-280.

Tobler, W.R. (1970), A computer model simulation of urban growth in the Detroit region, *Economic Geography,* 46(2), pp. 234-240.

Vázquez, Sergio Tezanos (2008), *The Spanish pattern of aid giving*, Working Paper 04/08, Instituto Complutense de Estudios Internacionales.

Ward, Michael and Kristian Gleditsch (2008), *Spatial regression models*, London: Sage.

2. Competing for Scarce Foreign Capital: Spatial Dependence in the Diffusion of Double Taxation Treaties

2.1. Introduction

Double taxation treaties (DTTs) are international contracts which are concluded between two independent tax jurisdictions and determine to what extent either of the contracting states is allowed to tax revenues of companies operating in both countries. The ultimate aim of such a treaty is therefore the division of the international tax base. Furthermore, there are a number of sub-motives. Even though the OECD (2005: 52) states that "the principal purpose of double taxation conventions is to promote, by eliminating international double taxation, exchanges of goods and services, and the movement of capital and persons", one of the most important objectives is the exchange of information on the revenues of companies in the respective contracting state to enforce domestic tax laws and combat tax evasion (Doernberg 2004). Apart from lessening administrative efforts as well as uncertainty faced by investors, curtailing tax evasion is also one of the reasons why DTTs harmonize calculation methods and definitions. This is particularly true for internal transfer pricing (Davies 2003). A motive that is especially important for developing countries is enshrining the effectiveness of tax related investment incentives to attract foreign capital, which can be neutralised without cooperation under certain conditions (Doernberg 2004). Besides, Baistrocchi (2008) argues that membership of the international tax treaty network based on the OECD model is associated with a minimisation of communication and enforcement costs as well as exclusive access to resources such as the multilateral advance pricing agreement. Finally, tax treaties regularly contain a dispute settlement mechanism. Some authors argue however that the benefits of entering a DTT may go beyond the mere treaty provisions in a way that signatory countries may gain "international economic recognition" (Dagan 2000: 32) or convey a "credible commitment to predictability and legal stability" (Baistrocchi 2008: 383). This argument might be even more relevant for developing countries as their regimes might be perceived to be less reliable and predictable (Arnold, Sasseville und Zolt 2002).

On the other hand, concluding a DTT is accompanied by a series of costs. First, the process of negotiating and ratifying the treaty ties up administrative resources.[7] These costs increase if the tax systems of the negotiation parties differ widely and if versions in different languages have to be synchronised. Second, sometimes the domestic tax laws have to be adapted to the provisions of the treaty. In this case the national fiscal sovereignty is curtailed. However, most important is the potential loss of tax revenues at least for one

[7] These costs can be quite significant; Anecdotally, Shelton (2004: 1.74) points out that the negotiation of the Netherlands-US treaty took more than ten years and consumed probably several person-years of work.

contracting state. Most DTTs favour residence over source taxation: the domicile country of the investor is allowed to tax the worldwide income while taxation in the host country (and thereby the source country of the taxable income) is confined to the withholding taxes whose upper limits are set in the treaty negotiations.[8] Due to the reciprocity of FDI flows, benefits offered to an investor from one contracting state should - in theory - be balanced out by the benefits given to the other country's investors in the own country. Especially between developing and developed countries, FDI flows (and stocks) however are often highly asymmetric, so that the revenue sacrifice is regularly one-sided with the net-capital importer (Easson 2000). Thereby, the loss of tax revenues for the net-capital importer is generally increasing in the degree of asymmetry and decreasing in rising withholding tax rates.[9] This problem is even more significant, as corporate taxes play a much more dominant role in funding public expenditures in developing countries as compared to richer nations, while the overall level of tax revenues (as a percentage of GDP) is considerably higher in the latter (Gordon and Li 2009).[10] After concluding a DTT in an asymmetric dyad, a larger share of the tax revenues accumulates in the net-capital exporting country. This can be illustrated by a simple example:[11] There are two countries, *E* and *I*, each with one investor that only invests in the other country. The *E*'s investor invests 100m in country *I*, while *I*'s investor only invests 20m in country *E*. Domestic corporate tax rates are 20 percent in country *E* and 30 percent in *I*, respectively. For reasons of simplicity, one would assume that without a tax treaty both countries avoid double taxation by exempting

---

[8] Repatriation of profits from a foreign subsidiary to the parent company usually works through dividends, interest, and royalty payments. Since these payments are costs to the subsidiary firm, they are not subject to taxation in the host country as part of the income of the subsidiary. Most host governments tax these repatriations through withholding taxes at least partly (Chisik and Davies 2004a). Ceilings for the withholding tax rates are agreed upon in tax treaties for these three types of income and usually lie in the range of 0 to 15 percent as proposed by the OECD model treaty. Furthermore, they are almost always identical for both treaty partners (Chisik and Davies 2004b). It is assumed throughout this chapter that domestic non-treaty withholding tax rates are equal to the domestic corporate tax rates, i.e. that there is no discrimination between foreign and domestic investors nor between repatriated and reinvested profits. The non-treaty withholding tax rates can be quite high, often around 25 to 30 percent and sometimes even higher (Easson 2004).

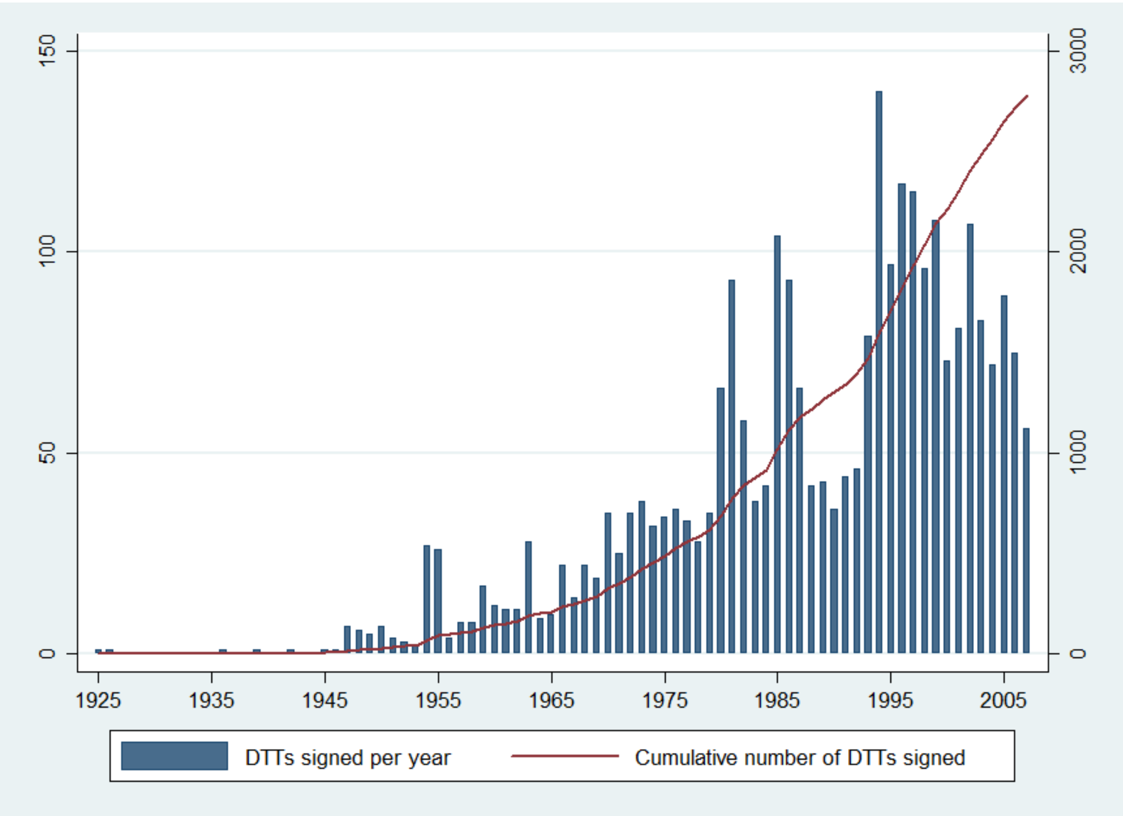
[9] Chisik and Davies (2004a) show theoretically that asymmetric countries will negotiate higher withholding tax rates. An empirical analysis of US treaties and treaties within the OECD confirms their finding. They additionally find evidence that highly asymmetric countries are less likely to enter such a treaty. This alleviates the problem to some extent. However, their first sample is restricted to OECD members and the second sample includes US treaty partners, but among them, there are only six developing countries (Indonesia, Egypt, Pakistan, China, Morocco, and India).

[10] Using data from between 1996 and 2001, Gordon and Li (2009: 857) show that while corporate income tax contributes to 19.3 percent of revenue in developing countries, compared with 9.7 percent in richer countries, the latter collect 25.0 percent of tax revenues as a fraction of GDP as opposed to 17.6 percent in the former.

[11] Recent research (Barthel et al. (2010) and Neumayer (2007)) suggests that the conclusion of a DTT on average leads to higher bilateral FDI stocks. This has two opposing impacts on the level of tax revenues in the host country: if the withholding tax rate is lower than the domestic corporate tax rate, the tax authority collects less from a single company, but on the other hand overall tax revenues may increase due to the increased tax base. Yet, these additional capital flows may even aggravate the asymmetry if the DTT leads to more inflows to the net-capital importing country than to the net-capital exporter.

foreign profits from domestic taxation.[12] Without a DTT, country *E* collects 4m of tax revenues and country *I* 30m. If both countries negotiate a typical DTT, it is agreed that the host country retains 10 percent withholding taxes while the residence country has the right to full taxation but deducts the withholding taxes paid in the other country from the domestic tax bill. Now country *E* collects tax revenues of 22m (2m withholding and 20m domestic tax) and country *I* 16m (10m withholding and 6m domestic tax). Crediting the withholding tax paid in the other country leaves a net-tax revenue of 12m for *E* and 14m in *I*. Compared to the situation without a tax treaty, *E* increased its revenues by 8m while *I* is worse off by 16m.

Figure 9: Newly concluded DTTs per annum (left hand scale) and total number of DTTs signed (right hand scale)



Data Source: IBFD (2009).

[12] The result holds true if profits are taxed both in the residence country of the investor and in the host country or if both countries unilaterally avoid double taxation using tax credits. In the former case, by avoiding double taxation, a DTT leads to a decrease in the total tax burden of both investors, and both countries lose tax revenues. But, importantly, the net capital-importer loses more than the net capital-exporter compared to the original situation. If the agreed maximum withholding tax rates equal or exceed the domestic corporate tax rate, the position of the host country does not change. However, as Chisik and Davies (2004a) argue, at least among US treaties, the agreed withholding tax rate is in no case greater than either country's non-treaty rate.

The fact that concluding a DTT is not unambiguously favourable for both contract partners has not prevented its spread. As Figure 9 illustrates, over the last eight decades, nearly 2,800 DTTs concerning the taxation of income and capital have been signed.[13] As of 2007, 188 tax authorities have at least signed one tax treaty with France and the UK leading the ranking with 123, and 118 respectively, treaties signed (Table 1). Since the number of high income countries is limited, it is not surprising that the more DTTs a specific country signed, the higher the share of treaties is where the partner country represents a developing nation. As can be seen from the table, the fraction of FDI outward stocks covered by a DTT is usually very high at least for those countries that signed many treaties. Most importantly, as can be seen in columns 6 and 7, asymmetric FDI stocks are clearly the rule rather than the exception. For example, on average, in 2004 the FDI outward stock of the UK in a contracting partner was 16.7 times the respective outward stock of this partner in the UK, while 42 partner countries report no FDI in the UK at all.[14] The average ratios are considerably smaller if weighted by the relative shares foreign countries account for as a percentage of overall FDI coming from a source country, but even then asymmetry is highly prevalent. In fact, only Australia has a mean weighted FDI asymmetry ratio of close to one, indicating that it's FDI in- and outward stocks are nearly balanced on average. Furthermore, developing countries usually have a FDI ratio of less than 1, which means that the average inward stock is higher than the respective outward stock.

[13] Apart from tax treaties that are related to the double taxation of income and capital, there are also tax treaties regarding estates, inheritances, gifts, shipping and air transport as well as social security. However, since DTTs concerning income and capital are by far the most prevalent and are closely linked to FDI they are the focus of the analysis. Furthermore, because the goal of this analysis is to scrutinise the diffusion of DTTs, all figures stated refer to the number of DTTs signed which might differ from the actual number of treaties in force as some contracts were terminated.

[14] FDI data for the year of treaty conclusion is not available for many treaties. Yet, the adverse effects on the fiscal situation of the net-capital importer persist as long as the treaty is in force.

Table 1: DTT details and FDI stock asymmetry of selected countries

Rank [a]	Country	Number of DTTs	Percentage with developing countries	Fraction of outward stocks covered by DTT [b]	Mean FDI asymmetry ratio (unweighted) [c]	Mean FDI asymmetry ratio (weighted) [d]	N [e]	Number of DTT partners without FDI stock in country
1	France	123	64.2%	99%	6.5	2.7	58	9
2	United Kingdom	118	60.2%	98%	16.7	7.5	70	42
3	Norway	107	63.6%	85%	1.1	0.4	75	54
4	Denmark	105	59.0%	93%	4.1	2.4	81	25
5	Sweden	102	62.7%	100%	6.6	2.2	63	45
7	Switzerland	96	60.4%	92%	5.0	3.4	76	48
9	China	93	51.6%	94%	0.7	0.5	75	51
11	Germany	91	60.4%	98%	20.6	9.8	68	26
12	India	86	55.8%	100%	0.8	0.2	67	48
12	Netherlands	86	55.8%	96%	47.1	25.4	59	10
14	Russian Federation	85	56.5%	84%	1.1	0.5	68	45
15	Romania	84	56.0%	100%	0.7	< 0.01	71	47
19	Malaysia	76	60.5%	83%	2.4	7.0	59	43
21	South Africa	73	52.1%	93%	1.3	5.6	54	33
24	United States	69	47.8%	92%	25.3	7.7	58	10
31	Singapore	61	45.9%	72%	9.8	71.3	47	29
33	Japan	60	56.7%	92%	71.9	82.4	47	18
57	Australia	42	40.5%	87%	3.7	1.2	27	9
89	Argentina	19	21.1%	71%	0.2	0.1	11	2

Notes: [a] Countries ranked according to total number of DTTs signed; [b] Year 2004 chosen for data availability; [c] Calculated as the FDI outward stock of country i in country j over the FDI stock of country j in country i: FDI_{ij}/ FDI_{ji}; [d] weighted by $FDI_{ij}/ \sum FDI_i$; [e] Number of countries for which information on outward stock available.

Since the prevention of twofold taxation is nowadays done unilaterally through tax exemptions and tax credits by almost all countries, tax treaties have lost much of their original intend (Easson 2000). Nevertheless, the number of DTTs signed proliferated over the last decades, increasingly including treaties between two countries with asymmetric FDI stocks, in which the FDI stock from one contracting partner in the other state considerably exceeds capital stocks in the reverse direction. This development is remarkable since the revenue sacrifice in such cases is one-sided with the net-capital-importer. Mostly developing countries find themselves in such a position. This chapter seeks to investigate why these countries voluntarily resort to entering tax treaties. It is argued that even if capital-importing countries would collectively be better off not signing DTTs, individually they still have an incentive to sign DTTs. If other focal countries have entered such treaties, then they are likely to lose out even more if they themselves do not sign DTTs. In other words, the proliferation of DTTs is driven by spatial dependence among capital-importing countries caught in a prisoners' dilemma. A country's propensity to enter a tax treaty therefore depends on whether other countries, which compete with the country for scarce foreign capital, have already signed a DTT with the same capital exporter.

This is the first analysis which tackles this question empirically and provides results based on a large country sample over a long period of time rather than on anecdotal evidence. The remainder of the work is structured as follows: Section 2 briefly reviews the relevant literature on policy diffusion and on tax treaties. Section 3 provides a more detailed descriptive analysis of the DTT network while section 4 presents some theoretical explanations for the popularity of these treaties in asymmetric dyads. It is argued that positive network effects play a role and that countries relying on fiscal incentives to attract foreign investment need DTTs to enshrine the effectiveness of their tax incentives. Furthermore, tax treaties help to overcome the dynamic inconsistency problem that may emerge in such cases. Finally, the situation of net-capital importing countries can be described as a prisoners' dilemma leading to an outcome which is not optimal for the players. Section 5 briefly describes the data and the methodology used. In Section 6, the main results are discussed whereas some tests for the robustness of the results are presented in Section 7. In Section 8, the empirical work is extended to an analysis of conditional spatial policy dependence. Finally, Section 9 concludes.

2.2. Review of Literature

This work builds on previous work in two so far separated areas of research: on the one hand, the growing field on spatial dependence in policy affairs and on the diffusion processes of policy measures and on the other hand the theoretical and empirical literature on tax treaties. Spatial dependence in the diffusion of policy has been analysed in a variety of areas, such as social policy (e.g., Jahn 2006, Brooks 2007) and trade and investment policy (e.g., Simmons and Elkins 2004). Also, spatial contagion in tax and fiscal policy has been widely examined (e.g., Hayashi and Boadway 2001, Hays 2003, 2009, Basinger and

Hallerberg 2004, Rork 2003, Swank 2006, Plümper et al. 2009, Plümper and Neumayer 2010).[15]

Empirical evidence on policy diffusion in a dyadic setting, however, is by far more limited. Manger (2006) analyses the spread of preferential trade agreements (PTAs) and finds positive evidence for spatial contagion, as countries in a fear of trade and investment diversion are more likely to sign a PTA if competitor countries do so. He argues that the main reason is regional competition for trade and foreign direct investment, as geographically close countries tend to produce similar goods and export these to similar markets. However, instead of modelling these similarities explicitly with appropriate weighting matrices, Manger only uses inverse distance as a link between potentially competing nations. The basic argument is, as it will be also argued in the case of DTTs below, that developing countries find themselves in a classical prisoners' dilemma situation and enter a preferential trade agreement because other countries do so as well. While the results suggest that countries are not motivated by the total number of PTAs in the international system, Manger finds evidence that countries are geared to agreement conclusions of geographically closer countries.

Elkins et al. (2006) examine the diffusion process of bilateral investment treaties (BITs). These treaties enhance the international reputation of a potential host government by providing a commitment for respecting contractual and property rights of foreign investors. However, just like DTTs, these treaties curtail national sovereignty. By the same token as Manger (2006) they argue that the spread of these treaties is driven by international competition for capital. Testing several hypotheses, they find empirical evidence for the competition hypothesis, i.e. that countries are more likely to sign BITs if their most important competitors recently entered such treaties. Furthermore, they find that coercion and learning play a crucial role in explaining BIT conclusions. Neumayer and Plümper (2010) readdress this issue and find different channels through which these treaties diffuse. They argue that it is not only the aggregate number of treaties signed by competitor countries that matters; rather, a specific capital importing country has a higher propensity to sign a BIT with a capital exporter if competing countries have entered a treaty with the same exporter. Furthermore, they find evidence that capital exporters are influenced by treaty conclusions of other exporters, too.

[15] Hochgatterer and Leibrecht (2009) provide a very comprehensive literature review on tax competition.

In research, DTTs have been the focus of both legal scholars and economists. The latter approach tax treaties either theoretically to explain their role in avoiding double taxation or empirically to analyse their effectiveness in spurring bilateral FDI. Generally, the law literature on international taxation is broad and deep.[16] However, the number of publications specifically dealing with tax treaties is more limited by nature. Irish (1974) discusses several early reasons for the emergence of a network of unfavourable tax-treaties between developing and developed countries: (1) Developing countries feel obligated to accept any provision to remove impediments to FDI contained in the domestic tax law of developed countries and to provide assurances of stability to foreign investors. (2) Capital-exporting nations are disinclined to give up revenues that would be lost through greater taxation in the host countries, also because they perceive this as a further financial transfer in addition to official development assistance. (3) Developed nations are in a superior bargaining position and are able to dictate the terms of a treaty. (4) Developing countries are not fully aware of the adverse nature of tax agreements with a bias for residence and have not yet recognised that alterations can be achieved in certain areas or with particular countries. A very sceptical view on the necessity of tax treaties to avoid double taxation is expressed by Dagan (2000), calling them a "myth" and arguing that they "serve much less heroic goals, such as easing bureaucratic hassles and coordinating tax terms between contracting countries, and much more cynical goals, particularly redistributing tax revenues from the poorer to the richer signatory countries" (Dagan 2000: 939). Rejecting this view, Brauner (2003) stresses the achievements of tax treaties in bringing tax authorities from many countries at the same table, fostering communication among them and advancing international tax harmonisation. Contrary to the argument made here, he sees their "enthusiasm to conclude as many treaties as possible with developed countries" (Brauner 2003: 308) as evidence that developing countries have immensely benefited from these treaties.[17] Focussing on a hypothetical DTT between the US and Ghana as an example of a treaty between an industrialised country and an LDC, Christians (2005) analyses the potential of such a treaty to foster FDI from the US to Ghana. Arguing that (1) the scope of such a treaty would be too narrow as many non-income taxes are not covered, (2) that double taxation is a disappearing phenomenon due to global tax-competition and the wide-spread availability of opportunities for tax evasion, (3) that the provisions in a DTT would not differ significantly from domestic law, (4) that non-tax issues such as inadequate

[16] Literature surveys are provided by Avi-Yonah (1996), Brauner (2003), and Ring (2007).
[17] In Section 4 it will be argued that developing countries find themselves in the situation which can be described as a prisoners' dilemma and that signing DTTs is not the optimal solution for them.

infrastructure dominate investors' decisions and (5) that a DTT does not deliver a significant signal of stability to potential investors, the author concludes that a tax treaty would offer few commercial benefits to investors which renders a significant impact on investment unlikely. Baistrocchi (2008) seeks to explain why a representative developing country has an incentive to sign a DTT with a developed country based on the OECD model treaty, even though these treaties are strongly biased towards residence taxation. First, the author argues that developing countries face a prisoners' dilemma and follow the OECD model for fear of driving FDI away to competing countries. Second, the properties of a network market help to shed light on the emergence of DTTs in asymmetric dyads, that is, specific positive network externalities such as minimisation of communication and enforcement costs or reputation advantages are only available to members of the network.[18]

In the realm of theoretical economics, researchers try to answer the question whether DTTs are a superior method to avoid double taxation than unilateral approaches and which method should be applied in order to reach this goal (Hamada 1966, Musgrave 1969, Hartman 1985, Bond and Samuelson 1989, Janeba 1995, Davies and Gresik 2003, Davies 2003a, Chisik and Davies 2004).[19] Empirical evidence regarding the question whether the conclusion of a tax treaty in fact increases bilateral foreign investment, deserves closer attention since attracting further capital is one of the major motives for developing countries to sign DTTs with capital exporting nations. Quantitative work can be divided in the use of monadic and dyadic datasets. Di Giovanni (2005) examines the impact of various macroeconomic and financial variables on cross-border M&A activities as a component of FDI over the period from 1990 to 1999, covering 193 countries in a monadic setting. He finds that a DTT is accompanied by increased cross-border acquisition activities. Neumayer (2007) estimates the effect of DTTs on FDI to developing countries, using both dyadic outbound FDI stocks from the US, as well as the total inbound FDI stocks of developing countries and the FDI inflows to developing countries as dependent variables. The former dataset encompasses data from 1970 to 2001 and 114 host countries; the latter dataset covers 120 host countries from 1970 on for the FDI flows and from 1980 on for the stocks, respectively. Neumayer finds a positive relationship both for the dyadic and the monadic

[18] See Section 4 for a more detailed presentation.

[19] There are three methods to tackle the problem of twofold taxation (Dagan 2000): (1) Tax exemption, where foreign income is not taken into account in the computation of taxable income in the home country, (2) tax credit, where the taxes paid abroad are credited against the tax liability in the home country, and (3) tax deduction, where taxes paid abroad are treated as costs and are deducted in the computation of the taxable income in the home country. In the latter case, double taxation is offset only partially, for which reason DTTs either prescribe tax exemption or tax credit.

data as well as both for middle and low income developing countries. Turning towards pure dyadic approaches, Davies (2003b) examines the impact of treaty renegotiations over the period 1966 to 2000 on both inbound and outbound US FDI and finds that DTT renegotiations had no effect on FDI.[20] Focusing on US inward and outward investment stocks, Blonigen and Davies (2004) examine the influence of a DTT conclusion on the size of bilateral FDI. Using a dataset of 88 partner countries over up to 20 years from 1980 through 1999, they conclude that DTTs have no positive effect on inward or outward FDI. Blonigen and Davies (2005) broaden their research by using OECD data on bilateral FDI stocks and flows covering 23 developed source countries over the period of 1982-1992. They generally find no robust results. Depending on the estimation methodology, they find either a positive or a negative relationship between the existence of a DTT and higher FDI stocks and flows. Egger et al. (2006) estimate the effect of tax treaties on bilateral outward FDI from OECD source countries over the period of 1985-2000 with a two-step selection model and find a negative effect of DTTs on FDI. Coupé, Orlova and Skiba (2008) concentrate their research on the influence of both BITs and DTTs on the FDI flows from OECD into transition economies, covering 17 source and 9 host economies over the period of 1990-2001. No consistent results are found as the sign and statistical significance of the estimated treaty coefficients depend largely on the estimator used. Finally, Barthel et al. (2010) use a dyadic dataset encompassing FDI stocks from 30 home to 105 host countries, out of which 10 and 84, respectively, are developing countries. Covering the years from 1978 to 2004, they find a positive and significant relationship between the existence of a DTT and the bilateral FDI stock.

2.3. Descriptive Analysis of DTT Diffusion

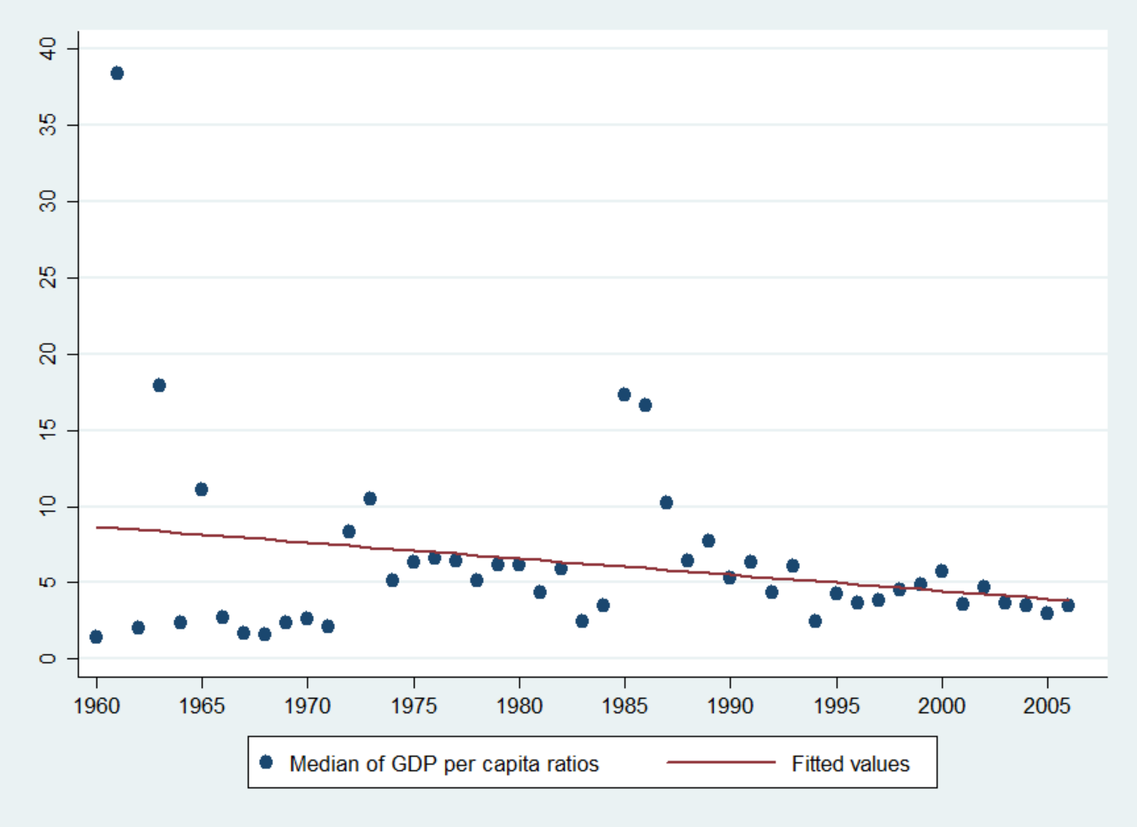
The agreement between France and Belgium closed in 1843, with its provision of information exchange regarding all documents and information which might improve the effective tax collection, can be seen as the forerunner of classic DTTs (Seligman 1928). The first bilateral treaties with the intention of avoiding double taxation were concluded between Prussia and Saxony and between Austria and Hungary as early as 1869. However, the Austria-Hungary and Prussia treaty agreed on 22 June 1899 is considered to be the first international double taxation treaty which resembles modern treaties. After World War I, concluding DTTs became more popular especially among countries on continental Europe

[20] Davies (2004) comes to similar conclusions.

(Easson 2000). The first model treaty was published in 1928 by a Group of Experts which had been convoked by the League of Nations in 1921 in order to develop possible solutions to the problem of international taxation. Even though the international tax legislation has become considerably more complex, the commentaries more extensive and some tax loopholes have had to be closed since then, this model treaty still forms the basis for all DTTs in force today (Graetz and O'Hear 1997).

So far, the diffusion of DTTs has not been examined empirically and also the amount of descriptive analysis is very limited. Easson (2000) distinguishes several geographical waves in the history of treaty formation: before the end of World War II, relatively few treaties were concluded and these mainly among countries in Europe. Until the 1970s, most DTTs were agreed upon between industrialised countries with market economies. Around 1970 several developing and less developed nations, mostly in Asia and Africa, started to conclude tax treaties with developed countries, with the goal of attracting more FDI. In the 1980s, some of the socialist countries began to enter DTTs, and during the 1990s, the transition economies and the newly independent states of the former Soviet Union followed suit. The latest wave encompasses the countries in Latin America and the Middle East. Kolm (2005) tracks the history of tax treaties from its early beginnings after the foundation of the German Empire and Austria-Hungary in the 19th century to the evolution of model treaties after World War II, while the focus is on the historical analysis.

Figure 10: Median of GDP per capita ratios of the two signatory states (in current USD)



Note: To avoid double counting, all ratios smaller than one are dropped.

As can be seen from Figure 9, until the end of World War II only very few treaties were signed. During the 1950s, an average of eleven treaties was signed per year. In the following decades, the diffusion of DTTs gained momentum with an average of 33 treaties p.a. in the 1970s, 65 treaties in the 1980s and 88 treaties per year in the 1990s. 1994 marked a peak with 140 newly signed treaties in this year alone. Since then, the extension of the treaty network lost some speed, but remains at an average level of 80 treaties per year since the millennium. Figure 10 illustrates the annual median of the GDP ratios of the two contracting states for the DTTs entered and a fitted trend line. There is no evidence that GDP asymmetry between the treaty partners changed over time. The trend line has a negative slope (-0.105), but the coefficient is not statistically significant (t-value: -1.18). Table 2 provides a geographical breakdown of 2,779 tax treaties which have been signed until the end of 2007.[21]

[21] Information on the contracting states, the type of treaty, and the date of signature are taken from IBFD (2009).

Table 2: Regional distribution of DTTs and percentage of possible combinations covered by a DTT

Region	No. of countries/ tax authorities [a]	Australia & Oceania	East Asia	Eastern Europe & Central Asia	Western Europe	Latin America & The Caribbean	Middle East & Northern Africa	North America	South Asia	Sub-Saharan Africa
Australia & Oceania	18	6 3.9%								
East Asia	19	28 8.2%	68 39.8%							
Eastern Europe & Central Asia	29	9 1.7%	152 27.6%	264 65.0%						
Western Europe	29	45 8.6%	161 29.2%	458 54.5%	206 50.7%					
Latin America & The Caribbean	43	4 0.5%	25 3.1%	23 1.8%	169 13.6%	68 7.5%				
Middle East & Northern Africa	20	1 0.3%	61 16.1%	129 22.2%	128 22.1%	8 0.9%	70 36.8%			
North America	4	5 6.9%	16 21.1%	46 39.7%	41 35.3%	20 11.6%	14 17.5%	1 16.7%		
South Asia	8	3 2.1%	38 25.0%	46 19.8%	57 24.6%	3 0.9%	32 20.0%	8 25.0%	21 75.0%	
Sub-Saharan Africa	48	2 0.2%	31 3.4%	34 2.4%	147 10.6%	6 0.3%	43 4.5%	11 5.7%	15 3.9%	56 5.0%
Total [b]	218	109	648	1,425	1,618	394	556	163	244	401

Notes: [a] Based on World Bank country list plus other tax authorities that signed at least one DTT; [b] To be read as number of treaty signatories from each region.

The two most active regions are Western Europe and Eastern Europe and Central Asia, hosting 1,618 and 1,425 signatories, respectively.[22] The principal diagonal shows treaties that were signed between two countries within the same region. In absolute terms, most inter-regional treaties were signed between countries in Western Europe and Eastern Europe and Central Asia (458), between Western Europe and Latin America and the Caribbean (169) and Western Europe and East Asia (161). When regarding the coverage of DTTs, expressed as the percentage of possible dyad-combinations covered by a tax treaty, an intra-regional bias is revealed. For instance, 75 percent of all possible country-combinations within South Asia are covered by a DTT, followed by 65 percent for

[22] Out of these, 206 treaties were signed within Western Europe and 264 within Eastern Europe and Central Asia, i.e. both contracting states came from the same region. For example, out of 1,618 signatories in Western Europe, 412 (2x206) represent intra-regional treaty conclusions. The remaining 1,206 represent DTTs signed with a partner country outside Western Europe (e.g., 169 with countries from Latin America and the Caribbean and 161 signed with East Asian nations).

countries in Eastern-Europe and Central Asia. Inter-regional coverage is the highest between Western and Eastern European nations (54.5 percent). DTT diffusion in Sub-Saharan Africa is limited both in absolute terms as well as in terms of coverage.

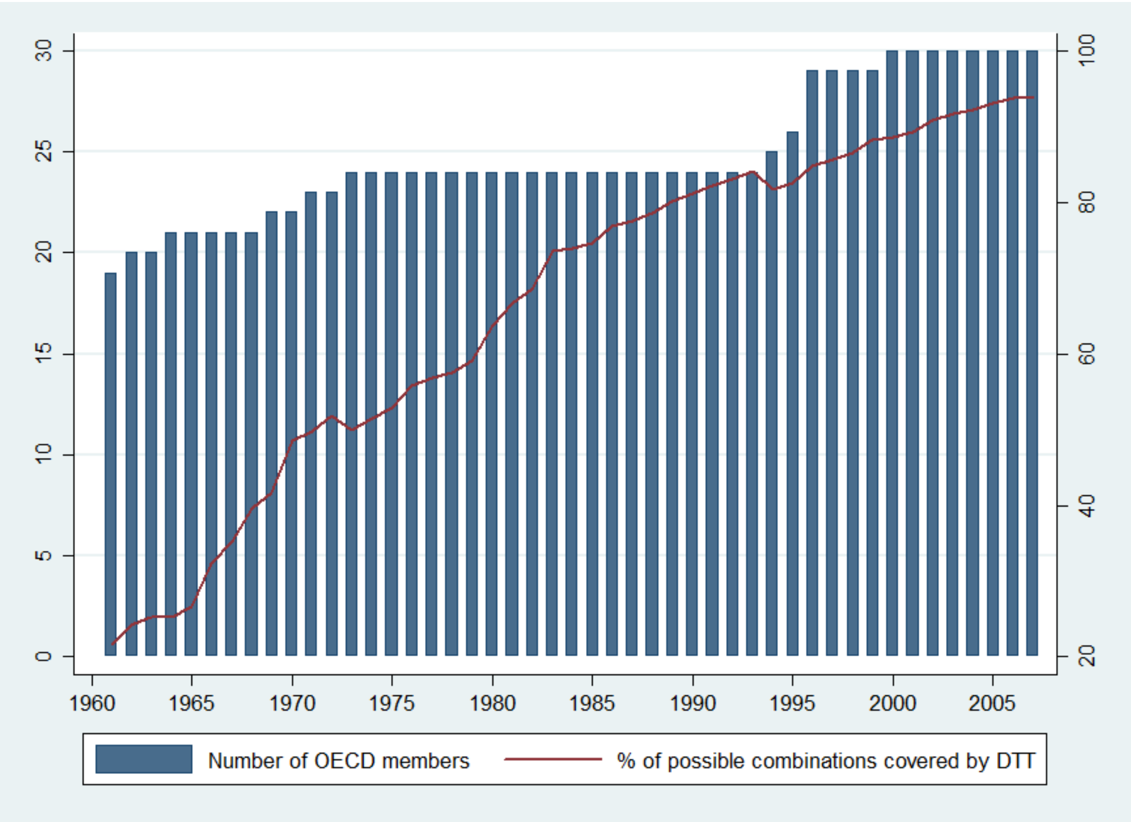
Table 3: Share of DTTs by income group dyad-type

Income group dyad	Share	Income group dyad	Share
High income - high income	22.1%	Upper middle income - upper middle income	4.9%
High income - upper middle income	21.5%	Lower middle income - lower middle income	4.5%
High income - lower middle income	20.5%	Lower middle income - low income	3.7%
Upper middle income - lower middle income	9.2%	Upper middle income - low income	3.3%
High income - low income	9.1%	Low income - low income	1.2%

Note: Classification according to World Bank GDP per capita limits.

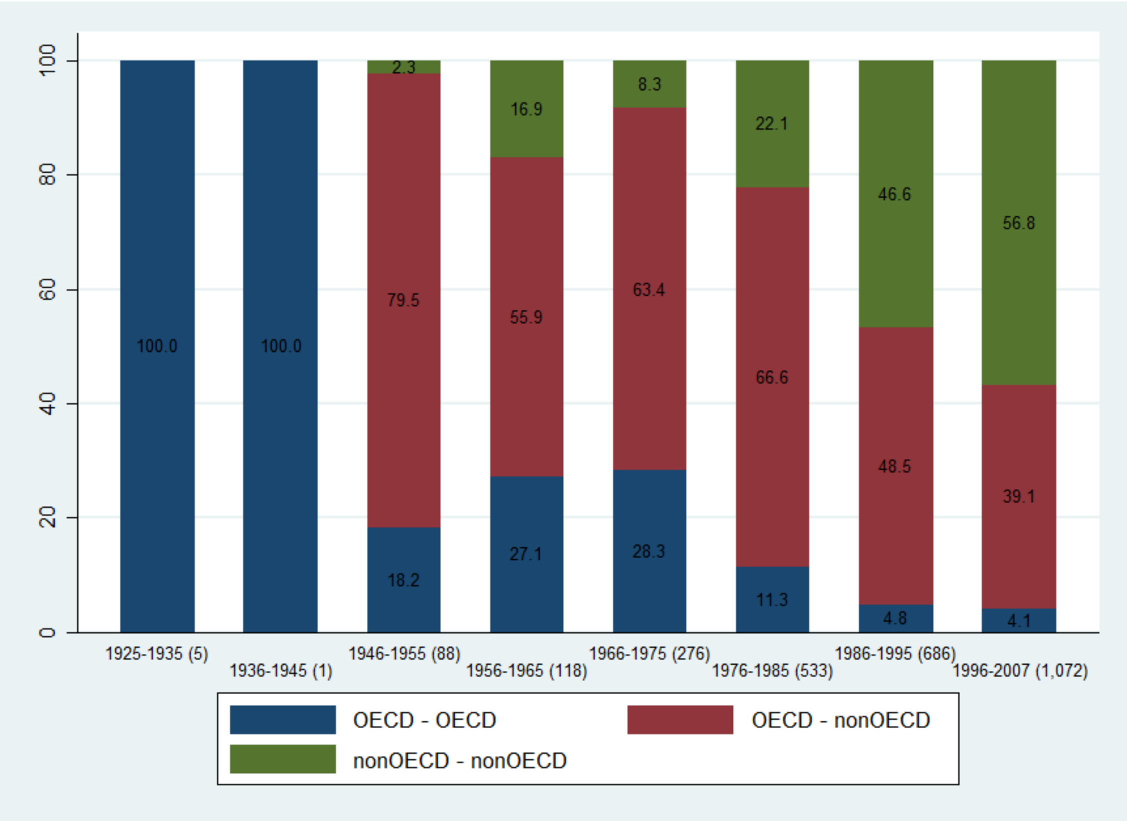
This pattern is confirmed when looking at the GDP per capita level of signatory states (Table 3). More than one fifth of all treaties where signed between two high income countries, and nearly three quarter comprise a high income country at least as one signing partner. Treaties between developing countries still represent only 26.8 percent of all DTTs signed with only 32 contracts between low income countries (1.2 percent).

Figure 11: Number of OECD members (left hand scale) and percentage of all possible combinations covered by a DTT (right hand scale)



A closer look at the OECD is justified as tax treaties are an important means for bilateral economic cooperation and because its members represent a more economically homogenous group of countries (thereby on average having less asymmetric FDI stocks), Figure 11 illustrates that the coverage has risen steadily since the foundation in 1961 reaching a level of 95 percent in 2007 despite an increasing number of member states.

Figure 12: Share of DTTs signed by OECD membership groups



Note: Based on actual OECD membership; Before 1961 foundation members taken as OECD members; Number in brackets behind the dates shows total number of treaties signed during the respective period.

The type of country that signed DTTs has changed considerably over time. Figure 12 pictures the share of tax treaties signed between OECD countries, between an OECD country and a non-OECD member and between two non-OECD countries. While the very few treaties before the end of the Second World War were signed exclusively among future OECD members, after 1945, already 80 percent of all treaties concluded in the following decade were signed between a prospective OECD member and a non-member. These mixed dyads accounted for two third of all signed treaties between mid 1980s, after that their relative importance decreased somewhat. This is due to a considerable increase in the number of DTTs signed between two non-OECD members. Their share started at 22.1 percent (118 out of 533 treaties) in the period from 1976 to 1985, and rose to 46.6 percent and 56.8 percent in the subsequent two decades. While treaties between developing (non

OECD) countries became more widespread over time, the share of newly signed DTTs between two OECD members declined steadily since the early 1970s, mirroring the rising total number of signed treaties as well as the increasing coverage of potential treaties within the OECD.

2.4. Theoretical Foundations

Putting aside negotiation costs and costs of curtailed fiscal sovereignty, tax treaties between two symmetric countries are beneficial for both. However, as shown in the introduction, DTTs among asymmetric countries, i.e. where one signatory state is a net-capital importer and the other one a net-capital exporter, are a frequent phenomenon. This is despite the fact that in such dyads one treaty-partner can loose significant amounts of tax revenues compared to a situation without a DTT.[23] This section tries to find answers to this seemingly irrational behaviour. One explanation might be that the other benefits associated with a DTT conclusion simply outweigh the costs and that the net-benefit therefore is positive. Without consideration of the cost of negotiation and ratification, the costs of lost tax revenues increase with the asymmetry of the FDI stocks and decrease with the level of agreed withholding taxes. On the other hand, some benefits, such as administrative savings and enforcement of domestic tax laws, decrease with the amount of outward FDI and are zero if one country has no outward investment in the partner country. In fact, as can be seen from the last column of Table 1, for each of the major DTT-signatories, there is a large number of partner countries that do not report any FDI outward stocks in this country. For example, out of 70 UK treaty-partner countries, for which FDI data is available, 42 have a FDI stock of zero in the UK. Another potential benefit is the increased FDI from the DTT partner. Barthel et al. (2010) show that a DTT, on average, is associated with an approximately 30 percent higher FDI stock and that this effect is larger for developing host countries. Here, two opposing effects occur: on the one hand, if the agreed withholding tax rate is lower than the domestic corporate tax rate, the host country will collect less from a single firm, but since there is more investment, the taxable profits should rise, too. A comprehensive analysis of the net-effect needs to be done on a country-level and needs to take into account not only the level of the applicable withholding and corporate tax rate, but

[23] If not stated otherwise, it is assumed that a DTT is based on the OECD model treaty, which has an explicit residence-bias in Article 21 and limits taxation in the host country to withholding taxes. For a justification of uniform treatment of all treaties, see Section 5. However, this problem is not confined to developing countries. In the aftermath of World War II, Western European countries agreed to enter DTTs with a strong emphasis on the resident principle to attract US capital and technology to rebuild and modernise their war-ravaged economies (Irish 1974).

also country-specific tax calculation rules and tax incentives. This is beyond the scope of this chapter. However, if a host country has no outward FDI in the partner country, if its domestic tax rate is 30 percent and if the inward FDI is increased by 30 percent, a withholding tax of 23 percent would be necessary to maintain the original amount of tax revenues. This is substantially above the conventional levels ranging from 0 to 15 percent. Given the ambiguous effect of a DTT on the overall tax revenues and the existence of positive effects from FDI that are hard to quantify, it is impossible to ascertain the net-effect of a DTT for a large sample of countries. This needs to be undertaken on a case study basis. However, it is clear that a net-capital importing country would be better off if it could attract the same amount of FDI without entering the kind of tax treaties that are common today. In the following, four other possible explanations for the widespread emergence of tax treaties in asymmetric dyads will be discussed: positive network effects, strategic interaction among host-countries, fiscal incentives, and the dynamic inconsistency problem, whereby the latter two are more extensions of the second rather than independent explanations.

Positive network effects

Network markets normally share three characteristics (Baistrocchi 2008: 359): (1) network externalities, (2) expectation and (3) a lock-in effect. Network externalities occur if the benefit for each member of a network increases in case the number of members rises. A classical example is the telephone market, where having a telephone is the more beneficial the more other people own one (Katz and Shapiro 1985). Expectation denotes the phenomenon that a certain standard may prevail over others, not necessarily because it is superior, but because it is promoted by an influential player. For example, the fact that the MS-DOS standard was initially promoted by IBM was a more important success factor than its technical superiority (Besen and Farrell 1994). Finally, the lock-in effect describes the situation that an established standard survives even it is inferior to a new arriving technology. An example is the QWERTY typewriter keyboard (David 1985). As illustrated by Baistrocchi (2008), the positive network externalities can deliver an explanation for the emergence of tax treaties in asymmetric FDI settings.[24] The argument is that the tax treaty

[24] Baistrocchi (2008) focuses on the OECD model as a basis for treaty negotiations. Since this model treaty favours residence over source taxation, it aggravates the problem of reduced tax revenues in asymmetric dyads. However, as argued in Section 5, because most existing treaties are based on this model treaty and the UN model treaty became more similar to the OECD model over time, this argument should be applicable to all treaties.

network produces positive network externalities only for its members, while the extent of benefits in some cases depends on the underlying model treaty. Communication costs, i.e. the cost of disseminating information about the content of the law among the members of society (which includes foreign investors) are a major cost block in legal institutions. Since most of the treaties are based on the same model and are usually provided in English, the communication costs are minimised within the tax treaty network. As the network grows larger, the average communication costs diminish (Baistrocchi 2008). The same effect of reduced average costs occurs with respect to enforcement cost, i.e. with applying the law to a particular dispute. For instance, domestic courts in developing countries gain access to legal sources (e.g., case law produced by foreign domestic courts interpreting OECD based tax-treaties) which are unavailable to countries outside the OECD network (Baistrocchi 2008). Furthermore, membership in a growing international network of tax treaties might radiate a stronger commitment to stability than a single bilateral DTT. If the cost-benefit analysis discussed above leads to a negative net-benefit for a specific country (with regard to a specific potential treaty partner), the cost-saving effects and increasing benefits of an extending DTT-network can change the calculation for certain dyads on the margin. Here, treaty conclusions by other countries increase the probability that other countries join the network by negotiating DTTs. As long as the benefits do not depend on the size of the outward FDI stock, this argument helps to explain the diffusion of tax treaties in symmetric as well as in asymmetric dyads. However, as Dagan (2000: 996) concludes, developing countries "have to sacrifice more to become a member of the 'treaty-club'".

Strategic interaction among host countries

Since FDI can provide a multitude of potential benefits, it is not surprising that host countries engage in a fierce competition for these long-term investment flows. One positive impact is the provision of new capital which allows additional investment in both physical and human capital. This is particularly beneficial for developing countries in which liquidity constraints are more prevalent (Busse and Groizard 2006). Moreover, foreign investment inflows are generally seen as a means to integrate new knowledge from abroad. The theory of the multinational firm states that multinational enterprises (MNEs) have a technological advantage over domestic firms that outweighs the cost associated with operating in external markets (Caves 2007, Markusen 2002). On the one hand, local companies may benefit from the influx of new knowledge through imitation and learning (Findlay 1978, Mansfield and Romeo 1980, Blomström 1986). On the other hand, increased competition in local markets, enhanced human capital mobility among firms (Fosfuri et al. 2001, Glass and Saggi 2002) and vertical linkages (Rodríguez-Clare 1996, Markusen and Venables 1999) serve as catalysers for knowledge diffusion. All these factors can raise the productivity level and promote a higher growth rate.[25] As discussed above, another advantage of increased foreign investment is the generation of taxable income in the host country (Gresik 2001). Finally, compared to other capital flows, FDI exhibits a greater continuity compared to short term capital flows in times of crisis.

A tax treaty is only one measure among a variety of factors that determine the attractiveness of a jurisdiction as a place to invest. Some of these factors, such as the geographical location of a country, the endowment with natural resources, the cultural proximity to major source countries or the access to natural transport infrastructure (e.g., rivers), are beyond the control of the host governments. Others, such as the education and productivity of the labour force, the quality of transport and TCT infrastructure or the purchasing power of the population, cannot be influenced within a short time horizon. However, there are some factors which can be altered by the host government and for which improvements can be achieved more rapidly. Unilateral examples for these encompass the a favourable national tax system (e.g., tax rates, simplicity, transparency, certainty and stability in the application of the tax law), fiscal incentives or alleviating red tape, while bilateral and multilateral measures can be the conclusion of bilateral investment treaties (BITs), regional trade agreements (RTAs) or double taxation treaties. A tax treaty can increase FDI by reducing uncertainty, by lowering the effective tax rate, by abating administrative efforts or simply because of positive signalling effects.[26]

However, as outlined above, a DTT is not unambiguously favourable for both treaty partners in asymmetric dyads, because the net-capital importer might face a loss in tax revenues which might not be offset by increased inward FDI nor by other positive effects of a DTT. Therefore, after entering a DTT with a developed country, a representative developing country might be worse off compared to a situation without a treaty. Yet, due to international competition for FDI, a developing country still can have an incentive to conclude such a tax treaty, because its own situation without a treaty deteriorates if competitor countries enter treaties with major capital exporters and thereby gain a

---

[25] Markusen (2002), Navaretti and Venables (2004), Helpman (2006), and Caves (2007) provide extensive literature surveys.

[26] However, on the other hand, a DTT might also lower FDI flows because the information exchange and transfer pricing provisions discourage tax-evading FDI on the margin.

competitive edge. As pointed out by Guzman (1998) in the comparable case of BITs, an individual country has an incentive to negotiate a treaty with potential investors to make itself a more attractive location relative to other potential hosts, but developing countries as a group presumably will not have any benefit.[27] If similar host countries act as a group, there would be less competition and more market power for 'sellers' (of investment locations). As a consequence they would be able to increase the price (in terms of less fiscal and non-fiscal investment incentives). Even though this would discourage FDI on the margin, the overall gains would outweigh the losses. Yet, this behaviour cannot be observed in practice. One possible explanation is that this situation can be described as a classical prisoners' dilemma.

In the following, India and China as two emerging countries competing for the same FDI are used as a case study, while a more comprehensive empirical analysis is presented in Section 6.[28] With 1.32bn people living in China and 1.12bn in India, respectively, both countries lead the table of the largest countries in terms of population. Yet, regarding economic size the differences are more pronounced – China produced a GDP of USD 3,382bn in 2007, compared to USD 1,177bn in India. This is equivalent to a GDP per capita of USD 2,562 in China and USD 1,050 in India, exhibiting considerable difference in terms of development level and purchasing power. Both countries are export orientated; however, China's exports of goods and services amount to 42.5 percent of GDP, twice as much as India's 21.2 percent. The same discrepancy is evident when comparing the FDI net inflows: while China attracted USD 138,41bn of foreign capital, or 4.1 percent of GDP, India only received USD 22.95bn or 2.0 percent of GDP (all data for 2007, World Bank 2009a). Locking at the Ease of Doing Business Indicators produced by the World Bank, China is ranked 89 out of 183 in the overall ranking and 125 in the paying taxes category. This is slightly higher than India on rank 133 and 169, respectively (World Bank 2009b). In the Corruption Perception Index published by Transparency International, China and India are listed on rank 79 and 84 in 2009, achieving a score of 3.6 and 3.4, respectively (Transparency International 2009). Taken together, even though there are some differences between the two countries, they might be regarded as substitutes from an investor's

[27] Guzman (1998) seeks to explain the apparent paradox that developing countries on the one hand fought to defeat the "Hull Rule" as customary international law which requires "prompt, adequate, and effective" compensation after expropriation, but on the other hand "enthusiastically signed BITs with developed countries" (Guzman, 1998: 666) which guarantees the investors even stronger protection of their investment against expropriation.

[28] Baistrocchi (2008) uses a law case involving China and India as a specific example of how both countries are worse off due to competition for FDI. However, this can be generalised to illustrate their respective strategic behaviour when entering treaty-negotiations as net-capital importers.

viewpoint. This is mirrored in an export market similarity of 0.62 in 2005 (down from 0.71 in the year 2000) and an export product similarity of 0.78 (down from 0.86 in 1994).[29] Both countries engaged heavily in the international DTT network, even though there are considerable differences in their starting points. India has signed 86 treaties in total, the first with Sierra Leone in 1956, followed by treaties with Sweden and Germany in 1958 and 1959. The first treaty with Japan in 1983 marked the beginning of comprehensive treaty negotiations in China resulting in treaties signed with 93 partner countries at the end of 2007. Both countries are on average net-capital importers with a weighted FDI asymmetry ratio of 0.5 in the case of China and 0.2 in the case of India. Their respective treaty networks overlap considerably, as 70 countries have signed a treaty with both India and China. For all treaties, the median time difference between the years of treaty signature with a partner country for both countries is five years, indicating that if one country entered a DTT with a specific contracting partner, the other country followed after five years. When focussing on treaties after 1983, when China first joined the international DTT network, the median time difference is reduced to four years. Before 1983, India has signed DTTs with ten OECD members, with eight of them China concluded a treaty within three years after its first appearance on the treaty stage. Regarding mutual treaty partners, since 1983, India was the first to sign a treaty with a specific country in 31 cases, while China was first only in 20 cases (in five cases, China and India concluded a DTT with a country in the same year). This suggests that India acted as a leader more frequently and China as a follower.

Formally, the respective positions of China and India can be displayed graphically (Figure 13). For illustrative purposes and to abstract from other FDI influencing factors, it is assumed that India and China are equally good at attracting (a specific type) of FDI. Furthermore, their decisions shall not be influenced by other parties, such as other competitor nations, so that both India and China only have to observe the behaviour of the other country. Both countries can choose between two alternative strategies: co-operate with each other and not to engage in harmful tax competition, or defect by concluding a tax treaty with major capital exporters.[30] Figure 13 shows the four possible combinations in such a situation and the respective outcomes. The numbers stand for the payoff for each outcome and are ordinally ranked, with 1 > 2 > 3 > 4 for each player. The payoffs to each

[29] See Section 5 for a detailed description of these measurements.
[30] These major capital exporters are basically OECD countries. Treaty conclusions with other nations should not induce the same pressure as they are less important in terms of exported capital. Furthermore, in these DTTs with non-OECD countries, asymmetry generally is less pronounced.

player for each potential solution are shown in the cells, with China's payoff being indicated by the first number, and India's by the second. Given that China does not conclude a treaty, India's best strategy (1) is to defect and to sign a DTT. Hence, the foreign investor will prefer India over China as an investment location. India's worst outcome (4) is to co-operate unilaterally, i.e. while China signs a tax treaty and attracts the investment, India does not. India's second best strategy (2) is to co-operate by not signing a treaty if China does the same. The third best strategy for India (3) is to sign a treaty if China also has a treaty. The same is true for China: the best outcome is to enter a treaty if India does not, followed by mutual cooperation and not to sign a DTT if India also refuses to conclude a DTT. The third best solution for China is to sign a treaty if India also has signed one, and the least preferable outcome is to co-operate unilaterally and not to sign a treaty, while India defects and signs a DTT. The dominant strategy for each country in this example is to defect and sign a treaty. Because both countries act in the same way, both will sign a treaty and thereby reach their third best solution. However, this is not the social optimum since both countries would be better off if they cooperated and mutually refused to sign a DTT.

Figure 13: Prisoners' dilemma among developing countries in the area of FDI

		India	
		Co-operate	Defect
China	Co-operate	2,2	4,1
	Defect	1,4	3,3

Source: Baistrocchi (2008).

Yet, in the resulting situation, where both countries sign a tax treaty, none is able to maintain a relative competitive advantage. In this case, both countries have an incentive to offer additional investment benefits to the potential investor in order to attract foreign capital as long as the gains of the investment to the country exceed the cost of these concessions. As a consequence, the two potential host countries might surpass each others in concessions (Guzman 1998). In a competitive market, this bidding up will continue until the net benefit enjoyed by the winner is zero. In this case, the firm does not have to share the surplus with the host country and it would choose the location which offers the highest overall return.[31] The country that wins the competition is able to attract further capital, yet little is gained from the victory as benefits in terms of employment, technology transfer, and tax revenues will be offset by incentives and concessions, such as reduced pollution control, tax breaks, or relaxed employment regulations. Baistrocchi (2008) suggests two basic solutions to overcome the prisoners' dilemma: in the contract solution, both countries enter a treaty in which they agree not to sign a DTT. However, an external authority is necessary to enforce this contract, which is not available. In this setting, each partner has an incentive to violate the contract and to sign a DTT in order to gain a (temporal) competitive advantage. Another solution emerges through iteration if the game is played repeatedly.

[31] This zero-benefit outcome will not be achieved if the two potential hosts differ. In this case, the bidding continues until one country drops out. The remaining country will have some gains (Guzman 1998).

Under this assumption, a country starts with a co-operative behaviour and defects only if the other partner has defected on the previous move. For this tit-for-tat solution three conditions need to be met. (1) The relationship between the players must be durable. This might not be given between relatively unstable developing countries, with a frequent turnover of both governments and officials.[32] (2) Co-operation needs to be based on reciprocity and (3) the players have the ability to observe defection when it occurs. In the context of tax treaties, these conditions are likely to be met, as the signing of a DTT is made public, even though the negotiation process is usually kept secret (Christians 2005) and as the non-defecting country could start treaty-negotiations itself. The widespread asymmetric DTT network in force shows that this tit-for-tat solution has not emerged in practice. Insufficient stability and predictability of the developing countries' behaviour can serve as one explanation; however, more significant is the fact that in practice there are more than two competing countries and collaboration is harder to achieve the larger the number of players involved. In a game with more than two players, each country has an incentive to sign a DTT to attract a competitive edge over its rivals and all other countries have an incentive to follow in order to equalise their competitive disadvantage. In a group of competing countries, the pressure on a single country to offset its deficit is the greater the larger the share of competitors that have already signed a treaty. The group of competing countries is defined as the set of jurisdictions which act as substitutes from an investor's point of view. In general, the degree of substitutability of countries increases with their similarity.[33] This strategic interaction among governments also helps to explain why nearly all DTTs are bi- rather than multilateral, as in a multilateral setting no single country is able to gain a competitive advantage. However, the situation might change as the net grows denser since then the choice is no longer between no DTT and a multilateral DTT, but between a bilateral and a multilateral DTT (Guzman 1998).

Apart from potential positive network effects and strategic interaction among governments, fiscal incentives and the lacking ability to make credible commitments to potential investors can also serve as a motive for developing countries to conclude treaties with asymmetric contracting partners. Rather than being completely separate arguments, however, they are closely linked to international competition for foreign capital, since DTTs help to maintain the effectiveness of investment promotion projects as well as to

---

[32] Axelrod (1984) shows empirically that tit-for-tat normally emerges if both players have a relatively stable identity.

[33] In the empirical analysis in Section 5, three different measures of similarity will be used: countries within the same region, countries that export to similar markets and countries which export a similar basket of goods.

mitigate the dynamic inconsistency problem – both factors might improve the position of a host country in the global rivalry for FDI.

Fiscal incentives to attract FDI

Tax incentives to attract inward foreign investment come in various forms and generally grant benefits in form of tax reductions that otherwise would be payable. Most commonly employed are reduced rates on corporate income taxes, property taxes or value added taxes (VAT), as well as tax holidays (i.e., reduction or exemption from tax for a limited period of time), investment credits or allowances, accelerated depreciation of capital assets, and reduced import duties (Easson 2004). In 2000, the UNCTAD conducted a cross-sectional analysis on the use of tax incentives to attract FDI in 53 countries, of which 44 are classified as developing countries. Tax exemption or tax holidays were offered in nearly all jurisdictions (96 percent), while investment allowances and exemption from duties and VAT are in place in three out of four countries (Table 4).

Table 4: The Use of Fiscal Incentives

Region	Countries in Study (developing countries)	Corporate Tax Rate [a]	Tax Exemption/ Tax Holiday	Reduced Tax Rate	Investment Allowance/ Tax Credit	Duty/ VAT Exemption/ Reduction	R&D Allowance
Africa	14 (14)	35-40	100%	29%	71%	79%	14%
Asia & The Pacific	17 (10)	30	94%	53%	65%	65%	59%
Europe & EIT [b]	10 (8)	30	90%	40%	70%	70%	10%
Latin America	12 (12)	30	100%	50%	100%	92%	8%
Total [c]	53 (44)		96%	44%	75%	76%	26%

Notes: [a] In percent of taxable income, regional averages; [b] Economies in transition; [c] Average weighted by countries in region. Data source: UNCTAD (2000).

In general, fiscal incentives are a widespread phenomenon; however, their effectiveness in inducing (additional) FDI inflows may be limited for two reasons:[34] (1) investors simply might not react to these incentives and (2) the tax-reducing effect of unilaterally offered

[34] See de Mooij and Ederveen (2003) for a meta-study on taxation and FDI. Buss (2001) provides a literature overview of tax incentives and firm location decisions. Reviewing the literature on the potency of tax incentives, Easson (2004: 68) concludes that their overall effectiveness is "somewhat mixed" and that "even though a majority of investors may be unimpressed by the availability of tax incentives there is abundant evidence that such incentives can be an important factor for certain types of investment."

lower tax rates can be offset in international taxation under certain circumstances. The latter effect occurs, if the home country of the investor (the source country of FDI), provides a tax credit for the amount of tax paid in the host country in order to unilaterally avoid double taxation, as it is common practice in almost all countries (Brooks 2008). Where no tax, or a reduced tax, is paid to the host jurisdiction because of a fiscal incentive, the investor pays the same amount of taxes as he would in a situation without a tax incentive, but pays a larger share of it in the home country. Stated differently, a tax incentive offered by a low-income host country has the effect of a revenue transfer to the high-income home country. If both states negotiate a so called tax sparing provision in their tax treaty, the home country agrees to credit the amount of taxes that would have been paid in the absence of a tax incentive, thereby preserving the incentivising effect of the reduced taxes and lowering the overall tax burden of the investor (Brooks 2008).[35] With the exception of the United States, tax-sparing provisions are usually included in DTTs between asymmetric dyads if the net-capital exporting (developed) country applies the credit method (Easson 2004). However, in recent years tax sparing has become somewhat less popular and several OECD members are now more restrictive in granting it in their tax treaties (Thuronyi 2003). Nevertheless, this provision can provide a motive to net-capital importers to sign a treaty with a major capital exporter, as it maintains the effectiveness of its tax commitments in the global competition for foreign capital.

Dynamic inconsistency problem

In a negotiation between the government of a potential host country and a major investor, the host country wants to encourage a major investor to invest while the company, on the other hand, wants to achieve the highest return possible and will only choose to invest if that country offers a higher anticipated profit than other potential locations. If the investor and the government were able to credibly commit themselves to an agreed set of conditions regarding the investment, the investor would select the most efficient location and a binding contract establishing the division of the surplus from the investment would be concluded (Guzman 1998). In this contract, the government could for instance pledge certain fiscal incentives or guarantee full repatriation of profits, while the investor commits itself to a certain level of employment, training of the workforce and technology transfer. However, due to a lacking superior authority, this contract between the investor and the

[35] This is a very simplified example to illustrate how tax sparing provisions work. See Brooks (2008) for a more comprehensive and critical view on the need for these provisions in DTTs.

host country is not enforceable. Furthermore, once the investment has been made, the host country does not need to attract the investment, but only has to treat the investor well enough to keep the investment (Guzman 1998). The distinction between a pre- and after-investment period is relevant, because the investment, once it has been established, cannot be withdrawn without cost to the investor. This irreversibility encompasses time and effort to dismantle or sell off assets, a reduced resale value, training of the local workforce, or adoption of the products to the local demand. Formally, $Z_t \geq \gamma Z_{t-1}$, where Z is the value of investment and $\gamma \in [0;1]$ the irreversibility parameter (Chisik and Davies 2004b). In a two period model, depending on the value of γ, the amount of investment that can be recovered in the year after the investment has been made (Z_t) is a share of the original sum of investment in the first year (Z_{t-1}). If $\gamma = 1$, the investment is completely irreversible, whereas the whole amount can be recovered if $\gamma = 0$. The host government can take advantage of this by extracting more value from the company as previously agreed, e.g. by increasing the tax rate.[36] Had the tax rate been set at this level during the negotiations, the investor might have chosen another – more profitable – location to invest. However, since it is not able to fully recover its initial investment, it might be cheaper for the company to pay the higher tax rates rather than to disinvest and reallocate the investment (Guzman 1998). Obviously, both the host government and the potential investors are aware of this dynamic inconsistency problem. As a consequence, the investor may decide not to invest. Yet, the host government wants to attract the investment and would, in order to receive the investment, be willing to bind itself to a set of commitments. However, as a credible commitment mechanism is not available, the investment may not take place.

Due to its focus on a single investment decision, this example neglects one important cost factor, namely reputational cost, which accrues to the host country if it breaches the original agreement and raises tax rates after the establishment has been set up. These costs of losing the reputation as a reliable host country occur if other potential investors observe the host country's behaviour and as a consequence eschew this country in their location decisions. As argued by Guzman (1998), however, even if the host country takes these costs into account, it will weigh the gains from breaching the agreement against any lost benefits caused by sanctions of other countries. There is no guarantee that balancing costs and benefits a priori completely abolishes the risk of time inconsistency. Working through

[36] Yet, if the resident country of the investors unilaterally grants tax credits to avoid double taxation, this increase in tax rate in the host country has no effect on the overall tax burden of the investor as long as the host's tax rate remains equal to or below the corporate tax rate in the residence country.

several channels, a bilateral tax treaty is one possibility to mitigate the dynamic inconsistency problem:[37] (1) DTTs clarify tax rules; (2) a tax treaty sets limits on certain host country tax rates; (3) due to the influence of tax treaty provisions on the domestic tax laws, a DTT lowers, though not eliminates, the risk of future alterations to existing tax laws (Dagan 2000). All three arguments work similarly, as a DTT reduces the opportunities for the host government to unilaterally change domestic tax laws in order to extract more value from the investment than previously agreed. This is particularly true for corporate tax rates, as maximum boundaries for withholding tax rates are specified for interest, royalties and dividends,[38] but also for the tax calculation methods which could provide a more subtle way to increase the fiscal benefits from an investment compared to raising tax rates. Terminating a treaty is more time consuming than the abrogation of domestic tax legislations which can be implemented over night in some jurisdictions (Baistrocchi 2008). Furthermore, the unilateral cancellation of a DTT could be perceived as a by far stronger negative signal than the positive signalling effect radiated by signing a treaty. This positive signal towards stability in investment and business climate and the dedication to protect and foster FDI which is expressed by the treaty partners in general can contribute to reduce the risk of adverse policy changes in the host country perceived by the investor. However, there are other methods, namely BITs, to provide an even stronger signal to stability than a DTT (Christian 2005).[39] A fourth channel through which a tax treaty can reduce dynamic inconsistencies in the host country's behaviour is the dispute resolution mechanism provided in most treaties. If an investor considers the action of the host government resulting in taxation not to be in accordance with the DTT, official representatives of its country may have to negotiate a reasonable solution with the tax authorities of the host country. Even though Article 25:3 of the OECD model treaty states, that the negotiating parties "shall endeavour to resolve by mutual agreement any difficulties or doubts arising as to the interpretation or application of the convention", there is no recourse if the authorities fail to reach a conclusion (Christians 2005). Finally, an argument brought

[37] Other mechanisms include: the setting up of a joint partnership with the host country (incentive for host to let the investor maximise profits), the placing of a few critical operations abroad (host gains little from expropriation), and the demanding of a signed agreement (non-compliance will increase international embarrassment to the host) (Guzman 1998).
[38] This argument is supported by Chisik and Davies (2004b: 116), pointing out that „without a treaty, governments cannot credibly commit to efficient tax rates, resulting in an inefficient equilibrium wiht high tax rates and low FDI. A tax treaty improves on this outcome by allowing governments to coordinate on a pair of Pareto-improving tax rates".
[39] As argued by Christians (2005), there is evidence that DTTs are not a sufficient signal of stability for US firms. For example, although Brazil and the US never concluded a DTT, there are significant US FDI flows to Brazil. On the other hand, despite a tax treaty between the US and Venezuela, FDI dropped due to concerns over regulatory and political stability in Venezuela.

forward by Elkins et al. (2006) for bilateral investment treaties, namely that these treaties increase the ex post costs of non-compliance, is also valid for DTTs. Tax treaties are negotiated between sovereign states and these state-to-state contracts implicate the direct interests of the resident country of the investor in a more direct way than contracts between an investor and the government. The home government has an interest in the broader principles of good-faith treaty observance. Any action by the host government that violates the DTT qualifies as a breach of the fundamental principle of international law that is '*pacta sunt servanda*' (treaties are to be observed). A non-compliance of an international treaty may have tremendous negative spill-over effects on other areas of foreign policy.

An illustrative example

The significance of DTTs in gaining a competitive edge against competitor countries, their role to maintain the effectiveness of fiscal investment incentive and their potential to overcome the dynamic inconsistency problem can be illustrated with a simple example:[40] Assuming a company is considering opening a branch in one of two potential host countries and that it will cost USD 60m for the initial investment and USD 10m for salaries, maintenance and so on. The costs and revenues can be regarded as discounted values and are therefore stated as lump sums. They are assumed to be the same for both host countries. Provided that the firm eventually decides to invest in one of the two locations, because it is offered a complete exemption from corporate taxes for five years and a reduced tax rate of 15 percent for the following 10 years. With these concessions, the investment yields expected revenues of USD 100m. The host country has agreed to the terms because even in absence of tax revenues the benefits in terms of employment and technology transfer leave a positive net-gain. Without the tax incentive, however, the host would not have been able to attract the investment. Under these assumptions, both the investor and the host country would be willing to sign a binding agreement in which they commit themselves to the terms stated above, as long as the agreement was credible and enforceable. However, due to the lack of a superior authority, the agreement is not enforceable, nor is it credible as the host has a positive incentive to change to conditions after the company invests. This is because the host country government knows that the investor cannot easily recoup its initial capital investment, i.e. that the investment is at least partly irreversible. Assuming now that the

[40] This example is a simplified version of the one presented in Guzman (1998). The author shows how BITs might help to mitigate the dynamic inconsistency problem, but ignores the problem of international taxation. Nevertheless, the point made regarding any non-tax investment incentive such as the preferential access to water and power, no local content requirement and the promise of unlimited profit repatriation remains valid.

host country, after the firm built its facility at a cost of USD 60m, imposes a tax of 40 percent on revenues, this leaves a after tax revenue of USD 60m. Clearly, if the firm had known beforehand that the host government would behave in that way, it would never have invested since the total cost of investment (USD 70m) is larger than the after-tax profit (USD 60m). Yet, because the company has already spent USD 60m, and supposing that the investment cannot be regained by tearing down or selling the facility, it must choose between continuing to operate the branch and face a loss of USD 10m, or closing down the operation and realise a loss of USD 60m. Obviously, the company will choose to remain in the host country.

The role of a tax treaty between the home country of the investor and the host country is threefold: (1) if only one of the two potential host countries has signed a DTT with the home country, this treaty could be the decisive factor on the margin since both countries are close substitutes. (2) If the home country eliminates twofold taxation unilaterally by granting tax credits for the taxes paid in the host country, as long as the DTT contains a tax sparing provision, the tax treaty secures preserving the effectiveness of the tax holiday and the reduced corporate tax rate. (3) Since the treaty specifies a maximum level of withholding taxes, it mitigates the dynamic inconsistency problem by limiting the amount of tax payment the host country can demand once the operation has been established. For all three reasons, the defeated host country also has an incentive to negotiate a DTT with the home country in order to offset its inferior bargaining position. This is true for the non-fiscal benefits as well, the defeated host country's tax incentives would be neutralised without a DTT or if it could not make a credible commitment.

Summing up: Why policy choices in other focal countries drive the diffusion of DTTs

With multiple competing countries, the pressure on a single country to sign DTTs is the greater the larger the share of competitors that have already signed a treaty, where competitors are simply other states acting as substitutes from an investor's point of view. In general, the degree of substitutability of countries increases with their structural similarity as a potential location for FDI. This strategic interaction among governments also helps to explain why nearly all DTTs are bi- rather than multilateral, as in a multilateral setting no single country is able to gain a competitive advantage.

Which other focal countries are countries likely to look at? Since competition among countries is the main hypothesized reason for the diffusion of DTTs, it is contended that countries look at their main competitors for scarce foreign capital as focal countries. Competitor countries can be identified in a multitude of ways, but in the empirical part three specific variables are chosen: one captures competition in the markets a country exports to, the second competition in the products a country exports to world markets, while the third captures the regional identity of countries.[41] Countries in the same region are competitors for scarce foreign capital as they are often functional equivalents for foreign investors, particularly for FDI of the market-seeking type, i.e. for FDI set up for the purpose of producing for a particular foreign macro-region. Countries competing in the same range of products and services are competitors for a specific type of investment in specific economic sectors producing these goods and services. Countries exporting to the same foreign markets are competitors in these foreign markets, but do not necessarily compete for FDI. This connectivity variable is included to examine whether countries are simply driven by the DTT signing activity of generally competing countries, rather than countries specifically competing for scarce foreign capital.

If competition for scarce foreign capital drives spatial dependence in the diffusion of DTTs, then the remaining question is why this does not lead to situation in which all possible country pairs are covered by a DTT? The reason is that while countries fear to lose out in the scramble for FDI if they fail to sign a specific DTT, they need to balance this cost against the costs of concluding the DTT. First, there is a post-treaty cost stemming from the potential loss in tax revenue already mentioned. Second, there is also an upfront cost. Despite the wide-spread use of model treaties, the negotiation of a tax treaty is a lengthy process which ties up a large amount of administrative resources. Especially developing countries lack the capacities to handle a significant number of simultaneous treaty negotiations. This is aggravated by the fact that some existing treaties are amended from time to time if general economic conditions change. Thuronyi (2010: 444) argues that "it would take a lifetime for most developing countries to negotiate a substantial number of treaties", since these countries only have weak capacity to administer treaties and their tax administrations are typically challenged in terms of human resources. However, the same

[41] While countries often look at their regional peers in terms of competition for FDI, competition is not the only reason why countries may regard their regional peers as focal countries. Countries often also learn from others in their region as well as imitate or emulate their behaviour. This is not a problem for the present analysis; rather it ensures that it also captures some other reasons for spatial dependence in the diffusion of DTTs.

restrictions apply, if less severely, to more advanced economies. The US Treasury Department has stated (US Senate Committee on Foreign Relations 2006):

> "The primary constraint on the size of our tax treaty network may be the complexity of the negotiations themselves. Ensuring that the various functions to be performed by tax treaties are all properly taken into account makes the negotiation process exacting and time consuming." (Statement of Patricia A. Brown, 02 February 2006).

As a consequence, policy makers have to prioritise potential partner countries with which they strive to negotiate a treaty. In line with the theoretical argument, the extent to which competing nations have already signed tax treaties will affect the relative cost of holding out and not signing a treaty. Moreover, industry structure as well as the composition and destination of exports of a country change over time. Consequently, also the set of competing countries is bound to change over time. As a result, more and more treaties will be signed over time, but the process of diffusion will slow down and potentially come close to a halt when the web of treaties covers the main competitors and concluding further DTTs generates higher costs than benefits to countries.

2.5. Data and Methodology

This section describes the data and methodology used to test the hypothesis derived in the preceding section, that the propensity for a specific dyad to conclude a tax treaty is influenced by DTTs signed between competitor countries, i.e. that there is spatial contagion between country pairs with respect to tax treaties. To begin with, the modelling of the dyadic setting is discussed, followed by a brief presentation of the dependent and the explanatory variables. Subsequently, the spatial weights used are discussed. Finally, the Cox proportional hazard model as primary estimation method is presented.

Choice of dyadic setting

Since a tax treaty is concluded between two states, a dyadic data setting where each observation consists of a country pair is the appropriate form of analysis. Basically, one can differentiate between directed and undirected dyads (Neumayer and Plümper 2010): while in the former case dyadic activity between the dyad members *i* and *j* initiates with *i* and is directed towards *j*, in the latter case such a distinction is either not possible or theoretically unimportant and $dyad_{ij}$ is therefore indistinguishable from $dyad_{ji}$. Examples of directed dyads generally encompass inter- and intraregional as well as international flows such as trade, FDI, remittances or migration, but also interstate armed conflicts where one state is an aggressor and another is the victim. On the other hand, examples for undirected dyads are bilateral treaties between more or less equal partners, for which the initiator of the negotiations cannot be identified. In the context of DTTs, modelling either a directed or an undirected dyad can theoretically be justified. Power based - or 'coercive' (Simmons et al. 2006: 790) - theories suggest that dominant capital-exporters such as the United States or Germany are able to control the agenda and start treaty negotiations according to their schedule and needs (Elkins et al. 2006). By the same token, these net-capital-exporting nations benefit most from a DTT in terms of tax revenues and enforcement of domestic tax laws, which gives them an incentive to approach potential treaty partners. On the other hand, if developing countries regard a DTT as an effective mean to attract more FDI, they could pursue treaties with targeted industrialised nations from which they expect the most potential investment. This argument is supported by Brauner (2003: 308), arguing that there are many cases in which developing countries wish to conclude tax treaties with developed countries, "which often reject their overture". Anecdotally, Nigeria sought to negotiate a DTT with the United States in 1978, however, despite lasting negotiations the treaty was never signed (Christians 2005). Furthermore, a major investor facing twofold taxation of its income can lobby both the home and the host government to sign a DTT. Since there is no conclusive theoretical argument for the choice of direction in a DTT-dyad and information on who started the negotiations is not available in individual cases, the present analysis is based on undirected dyads.[42] However, in the robustness section, following the lead of Elkins et al. (2006) and Neumayer and Plümper (2010) DTTs are modelled as directed from OECD to non-OECD countries.

Data description

The dataset covers 186 countries and spans 37 years from 1969 to 2005. Modelled as undirected dyads, 186 countries can be combined to 17,205 country-pairs.[43] The dataset encompasses 2,325 tax treaties, covering 87.8 percent of all treaties up to 2005 in the IBFD database. 270 treaties are lost as their date of signature is before 1969; another 53 treaties

[42] As a consequence, in the context of this analysis, $dyad_{ij}$ is the same as $dyad_{ji}$. Therefore, the dataset is halved to avoid double counting of identical observations. Furthermore, since a country cannot enter treaty negotiations with itself, all dyads $dyad_{ii}$ are dropped.

[43] The number of possible combinations can be calculated as ((186*186)-186)/2.

are excluded because either one or both of the contracting states is not in the sample due to a lack of data. With the exception of West Bank and Gaza and Greenland, these countries are small island states or mini-states such as Monaco and San Marino. Their exclusion should not affect the results.[44]

To begin with, the dependent variable is the signature of a DTT concerning the taxation of income and capital in a given year between a specific country-pair, which is measured by a simple dummy variable taking the value of one, if a dyad signed a DTT in certain year and zero if otherwise. Since the focus of the analysis is on the diffusion of tax treaties, the date of signature, rather than the date of ratification or the date of effectiveness is taken. For the same reason, all renegotiated contracts which replace a former treaty are excluded. Since the treaty partners can phrase their own wording of the agreement there is some heterogeneity among existing DTTs, however, most treaties follow the recommendations of the different model tax treaties, which are used as a basis and are adopted during the negotiation process. Several arguments justify the homogenous treatment of all tax treaties: First, the focus of analysis lies on treaties on the taxation of capital and income, which form a more homogenous subgroup of all possible tax treaties. Second, Avi-Yonah (2009) estimates that about 75 percent of the words of any arbitrary DTT are identical with the words of any other DTT. Third, the OECD Model Treaty is by far the most prevalent basis for existing DTTs. For instance, the pro-host country taxation Andean Model treaty is practically irrelevant since it has never been used as a basis for negotiations between a developed and developing country (Baistrocchi 2008). The UN Model Treaty, on the other hand, may be the preferred choice of developing countries when starting treaty negotiations.[45] Yet, they often lack the political clout to prevail in treaty negotiations with developing countries. Regarding the UN Model and the OECD Model as the two opposing ends of a continuum, however, there is some evidence that the larger a developing country, the closer the agreed treaty is to the UN Model. For example, the asymmetric tax treaty network of Brazil is closer to the UN Model, than the asymmetric DTT network of Kazakhstan (Baistrocchi 2008). On the other hand, in the 2001 update of the UN Model several changes were made "to bring the UN Model more in line with the OECD Model" (Kosters 2004: 7). Taken together, there is sufficient reasoning for a uniform treatment of

[44] Many of the excluded island states are classified as offshore financial centres, which can lead to a sample selection bias for this variable. However, the diffusion of DTTs among OFCs is not the focus of analysis.
[45] The UN Model Treaty was first published in 1980, motivated by the fact that the OECD Model does not adequately take into account the fiscal interests of developing countries. Compared to the OECD Model, the UN Model contains 27 adoptions with the aim of enlarging taxation in the host country vis-à-vis the residence country (Wijnen 1998).

DTTs and thereby to follow the example of almost all other empirical studies examining tax treaties.

Contrary to datasets consisting of directed dyads, where one country clearly is the sender (e.g., exporter) and the other country the recipient (e.g., importer), such a distinction is not possible in undirected dyads. As a consequence, because $dyad_{ij}$ cannot be distinguished from $dyad_{ji}$, it is not possible to easily include characteristics of only one country (e.g., GDP per capita); rather the relationship between the characteristics of both dyad members has to be modelled and theoretically justified. An example for such a linkage is a substitutive, represented by the sum of a characteristic of both countries (e.g. sum of GDP per capita of both dyad members) or a complementary, modelled through the product of a characteristic of both countries (e.g., product of GDP per capita of both dyad members), relationship. Another possibility is to take either the maximum or the minimum of the values of both countries, e.g. the higher value of GDP per capita. The same is true for directed linkages, such as trade or migration flows. The inclusion of relational variables (e.g., distance or bilateral contracts), on the other hand, is straightforward.

Besides the spatial lags, described in detail in the next subsection, the standard estimation model contains the following explanatory variables:

- *Product of populations*: The population of a country is taken as a measure of its size, while the use of the product implements a complementary relationship. Ceteris paribus, it is expected that two large countries are more likely to sign a DTT than a large and a small country or than two small countries, because more populous nations generally are more important players on the international stage.
- *Product of GDPs per capita*: To control for economic development, the product of the GDP per capita for both dyad members is included. As with population, it is anticipated that two rich countries have a higher propensity to enter a tax treaty than a developed and a less developed nation and particularly than two developing countries.
- *Bilateral Trade*: The sum of bilateral exports and imports is taken as an approximation for the intensity of the bilateral economic ties. To reduce the skewness of the data, the log is taken and to mitigate the potential endogeneity bias,

the variable is lagged by one year.[46] A positive influence on the propensity to sign a DTT is expected.

- *Product of both countries' openness to trade*: Openness to trade is defined as the sum of imports and exports divided by GDP. The product of both countries' openness to trade is taken to model a complementary relationship. Since two open countries should be more likely to sign a DTT, a positive sign is anticipated.
- *BIT*: A dummy variable taking the value of unity if the two countries have signed a BIT. These investment protection treaties are an important mean for bilateral economic cooperation. In order to secure cross-border investment, both countries commit themselves to a fair and equitable treatment of investments, to refrain from unreasonable and discriminatory measures as well as expropriation and to guarantee the free transfer of assets (Guzman 1998). Furthermore, a dispute settlement mechanism is included in most treaties. With respect to the positive signalling effect to potential investors, a DTT and a BIT might act as substitutes; however, since they differ in their key aspects, a complementary relationship and therefore a positive influence of a BIT is expected.
- *RTA*: Another mean for economic cooperation is the conclusion of a Regional Trade Agreement (RTA). This dummy takes the value of one if both signatory states are members of the same RTA. A positive association with the propensity to sign a DTT is expected.
- *OFC*: A dummy which takes the value of one, if one of the dyad members is classified as a tax haven or as an Other Financial Centre using the OECD country list. The OECD defines countries as tax havens, if a country (1) imposes no or low nominal taxes on the relevant income, (2) refuses to effectively exchange information with other jurisdictions, (3) lacks transparency, and (4) has no requirement that an activity must be substantial (OECD 1998).[47] The OECD advises its members not to enter, to limit or to terminate tax treaties with these countries (OECD 1998, OECD 2004). Therefore, a lower propensity to sign a DTT is anticipated if one of the dyad members is on the OECD list.
- *Diplomatic representation*: A dummy equal to one, if one country has a diplomatic representation in the other country. Since diplomatic relations are a signal for

[46] To keep zeros in the sample, the smallest existing positive value is added to each observation before taking the log.
[47] The absence of a requirement that an activity has to be substantial indicates that a country attempts to attract purely tax-driven FDI, i.e. an activity with little value added in the country such as booking centres (OECD 1998).

political cooperation and reduce transaction cost in the negotiation process, a positive influence is expected.

- *Distance*: It measures the distance between the capitals of both dyad members. On the one hand, a high distance increases transaction costs in the negotiation of a DTT; on the other hand, as far as a tax treaty provides a positive signal towards stability and investor-friendliness, these effects might be more important for remoter countries. The sign is therefore uncertain.
- *Product of political constraints*: Political constraints on the executive branch as an approximate measure for the institutional development of a county is included, as poor institutions could hamper DTT negotiations (e.g., by increasing the transaction cost) or as less discretion is supposed to render credible commitments to foreign investors more likely. On the other hand, in its role of providing stability for investors, a tax treaty might be more beneficial with countries with poorly developed institutions. Therefore, the sign is unclear. As these arguments apply to both dyad members, a complementary relationship is assumed.
- *Minimum years of independence*: Measures the years of independence of the dyad member, which became politically independent at a later stage. Since only independent countries can sign DTTs, newly independent countries might need to catch-up with other nations and sign treaties with the most important capital exporters. Therefore, the probability to enter a DTT is expected to decrease with years of independence.
- *Maximum number of DTTs*: Measures the higher of the number of existing treaties of either of the two countries in a dyad. As a country with a high number of exisiting treaties is likely to have signed DTTs with all important partner countries, the probability to negotiate a further DTT is anticipated to decrease with increasing number of existing treaties.
- *Country group dummies*: To control for inherent differences in the probabilities of signing a DTT between OECD and non-OECD countries,[48] a dummy taking the value of unity if the dyad consists of two OECD countries and a dummy that is equal to one if an OECD and a non-OECD country form the country pair is included. The omitted reference category are dyads in which neither of the two countries are member of the OECD.

---

[48] Only Western developed OECD members are taken as OECD members: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Portugal, Norway, Spain, Sweden, Switzerland, United Kingdom, United States.

- *Cumulative number of DTTs*: Two variables capturing the cumulative number of DTTs that country *i* and country *j*, respectively, have signed until year *t-1* to control for the general, but time-varying propensity of a country to enter such treaties.

Ideally, one would control for the size of bilateral FDI stocks as a measurement of foreign investment exposure in the partner country and the degree of asymmetry of the dyad. Unfortunately, this data is not available before 1970 and very scarcely until the early 1990ies. Furthermore, data availability is non-random with information more completely obtainable for OECD countries. Including this variable may therefore lead to a severe sample selection bias. Furthermore, as predicted by the gravity-model, empirical work has found a high explanatory power of market size and distance on FDI, both included in the model (Blonigen et al. 2007).

Spatial lag variables

To analyse the potential influence of DTT conclusions by competitor countries, a spatial lag model is used. Contrary to a spatial-x model, where the dependent variable is regressed on the weighted values of one or more explanatory variables in other dyads, and to the spatial error model, in which spatial dependence is modelled as a part of the error term, spatial lag models include the weighted values of the dependent variable in all other dyads as an independent variable (Anselin 1988). In undirected dyads as opposed to directed dyads, the flexibility to model spatial dependence is limited considerably. While in the former case, policy choices of a dyad can only depend on the choices of other dyads, directed dyads allow testing far more elaborated channels of policy diffusion (Neumayer and Plümper 2010). Ignoring other control variables, the basic model takes the form

$$y_{ij} = \rho \sum_{km \neq ij} w_{pq} y_{km} + \varepsilon_{ij} \qquad (7)$$

where y_{km} represents the total sum of DTTs signed by all other dyads and w_{pq} is the weighting matrix measuring the connectivity between $dyad_{ij}$ and other dyads *km* (Neumayer and Plümper 2010). Other than in a monadic data setting, where the weighting matrix simply represents the connectivity between unit *i* and unit *k* (e.g., reversed distance),[49] modelling the weighting matrix in a setting where dyads depend on dyads is

[49] The spatial weight itself can be directed (e.g., exports) or undirected (e.g., distance). Since all weights used are undirected, this discussion is set aside here.

more complex. Since each dyad member i and j may be influenced by the sum of DTTs signed by every other dyads except the $dyad_{ij}$, the weighting matrix w_{pq} contains the linkages between country i and all other dyads as well as linkages between country j and all other dyads. These linkages might be considered as complements or substitutes, i.e. in the first case both dyad members must be spatially influenced while in the latter it is sufficient that only one dyad member must be subject to contagion.[50]

To model the strength of connectivity between two units, three concepts of similarity are employed:[51]

- *Common region*: Dichotomous weighting matrix that takes the value of 1 if both the influenced and the influencing country are in the same region. This means that each dyad member is only influenced by DTT conclusions of jurisdictions in the same region. For instance, in the dyad Germany-Ghana, Germany is influenced by other European countries, while Ghana is influenced by all other nations in Sub-Saharan Africa. A complementary relationship between the two linkages in the weighting matrix is assumed, i.e. the dyad Germany-Ghana together is only influenced by DTTs signed between European countries (excl. Germany) with Sub-Saharan nations (excl. Ghana).
  Formally, this spatial lag is modelled as follows:
  $\sum_{km \neq ij} (w_{ik} \times w_{jm}) y_{km}$, with $i \neq k$ and $j \neq m$, where w_{ik} takes the value of one if both country i and k are within the same region (analogous for w_{jm}).
- *Export market similarity*: Two countries are supposed to be close competitors if they export to similar partners. Here, a substitutive relation between the two linkages is taken. In the Germany-Ghana example, a strong export similarity between Germany and third country can offset a low similarity between Ghana and a fourth country. On the other hand, the higher both export market similarities, the higher their overall weight in the weighting matrix. More specifically, in 2005, the export market similarity of Germany and Italy was 0.72 and 0.14 for Ghana and El Salvador, respectively. A DTT between Italy and El Salvador is therefore weighted with 0.86 (compared to 0.10 if a complementary relationship was assumed).

[50] Substitutes are modelled by taking the sum of the both weighting functions, while for a complementary relation their product is taken (Neumayer and Plümper 2010).
[51] The spatial lags were created using the ado-file *spundir.ado* by Neumayer and Plümper, described in Neumayer and Plümper (2009).

- *Export product similarity*: Two nations are supposed to be close substitutes as an investment location from the point of view of a foreign investor if they export a similar basket of goods, based on thirteen product categories. As with export market similarity, the linkages in the weighting matrix are assumed to be substitutes. Formally, these spatial lags are represented by: $\sum_{km \neq ij}(w_{ik} + w_{jm})y_{km}$.

To calculate the export market and export product similarity, an approach suggested by Finger and Kreinin (1979) is adopted:

$$Similarity(ab_t) = \{\sum_c Min[X_c(ac_t), X_c(bc_t)]\} \quad (8)$$

Where *a* and *b* are two countries exporting either a commodity *c* or to a market *c* and $X_c(ac)$ is the share of exports in commodity *c* or to market *c* of the total exports of *a* in year *t*. The similarity of *a* and *b* is the sum of the minima of the shares of a certain commodity (or a target market) of the total exports of *a* and *b*, respectively. The resulting index ranges from 0 to 1 and takes the value of 0 if the two countries export completely different products (e.g., country *a* only exports agricultural products and country *b* only ores and metals) and the value of 1 if both countries export exactly the same basket of goods (e.g., 50 percent of country *a's* and country *b's* exports are agricultural products and 50 percent ores and metals).[52] The resulting weighting matrix is row-standardised, i.e. for each row of the matrix, each cell is divided by its own row sum.[53] The spatial lag is then the weighted average of the number of DTTs concluded by competitor countries. Apart from having some convenient characteristics such as an intuitive interpretation of the coefficients, there is a theoretical reason behind row-standardisation: every country should be subject to the same "amount of contagion", regardless how many competitors it has. For example, in the case of common region, the sum of all contagious influences is the same whether it is a small or a large region. If the weighting matrix was not row-standardised, any result could be driven by the fact that regions differ in size. Furthermore, it could be argued that a

[52] Missing data in the export structure (e.g., country *a's* export in ores and metals) are implicitly set to zero in this approach, since the export shares of the remaining twelve product groups are calculated without ores and metals. As the minimum of the share of a certain product category for country *a* and *b* is taken, in this example country *a's* zero is smaller than any value for country *b* in this product group. As a consequence, if missing values are in fact positive and not zero values, the similarity measure underestimates the true similarity. Therefore, these similarity measures should be regarded as the lower bound. Contrary, simply taking the correlation between country *a's* and country *b's* export structure would overestimate the similarity in case of missing values.

[53] This is a very common procedure in spatial literature; yet, Plümper and Neumayer (2010) argue that since row-standardising changes the relative weights of influencing units, it needs to be well justified.

specific country observes the actions of countries in the same region more closely if the region is small. For instance, South Asia consists of eight nations and Sub-Saharan Africa of 48. India therefore should monitor the action of each of its seven regional competitors more closely and react adequately than Ghana with its 47 regional competitors. By the same token, the total influence of other dyads in terms of export product and export market similarity is taken as fixed with the weights depending on the degree of similarity where more similar dyads are assigned more weight than less similar dyads. Table 5 provides summary statistics and data sources for all variables.

Table 5: Summary statistics and data sources

Variable	Obs.	Mean	Std. Dev.	Min	Max	Source
DTT	212,244	0.007	0.081	0	1	IBFD (2009)
Product of populations (ln)	212,244	31.592	2.363	22.276	41.543	World Bank (2009)
Product of GDPs per capita (ln)	212,244	14.765	2.079	9.374	21.072	World Bank (2009)
Bilateral trade (ln, t-1)	212,244	9.986	7.765	0	25.237	Rose (2009)
Product of openness' to trade	212,244	5,296.2	4,604	65.432	94,121.9	World Bank (2009)
BIT	212,244	0.046	0.209	0	1	UNCTAD (2007)
RTA	212,244	0.079	0.269	0	1	WTO (2009)
OFC	212,244	0.209	0.407	0	1	OECD (2009)
Diplomatic representation	212,244	0.307	0.461	0	1	Bayer (2006)
Distance (ln)	212,244	8.782	0.705	4.54	9.90	Bennett and Stam (2005)
Product of Political Constraints	212,244	0.125	0.191	0	0.786	Henisz (2000)
OECD-OECD dyad	212,244	0.009	0.009	0	1	
OECD-nonOECD dyad	212,244	0.262	0.440	0	1	
Min. years of independence	212,244	36.718	17.10	2	81	CIA (2010)
Max. number of DTT (t-1)	212,244	24.428	23.98	0	118	IBFD (2009)
Cumulative number of DTTs country *i* (t-1)	212,244	12.993	19.52	0	120	IBFD (2009)
Cumulative number of DTTs country *j* (t-1)	212,244	15.987	21.73	0	120	IBFD (2009)
Spatial lags:						
W: Common region (product) (t-1)	212,244	0.049	0.081	0	1	own calculations, based on IBFD (2009) and World Bank (2009)
W: Export market similarity (sum) (t-1)	212,244	0.094	0.046	0.015	0.219	own calculations, based on IBFD (2009) and World Bank (2009)
W: Export product similarity (sum) (t-1)	212,244	0.100	0.050	0.012	0.279	own calculations, based on IBFD (2009) and World Bank (2009)

Note: W denotes the weighting matrix used to create the spatial lags. The spatially lagged variable is a dummy taking the value of one in years a DTT has been signed and in all subsequent years.

Estimation methodology

The aim of this analysis is to find out whether specific factors, particularly the treaty conclusions of competitor countries, have a significant influence on the propensity to sign a treaty for a specific dyad, i.e. how these factors influence the time until two specific countries conclude a DTT. Using this time in years as dependent variable, the influence of the spatially lagged treaties of competitor countries could be estimated by OLS. However, this estimation technique has several limitations in this setting (Box-Steffensmeier and Jones 2004): (1) duration data, i.e. years until treaty conclusion, are always possible and some observations have very long duration times which leads to considerable skewness in the data. Taking logs can mitigate, but not solve the problem. (2) Many dyads have not concluded a treaty until the end of 2007 and it is unknown when or whether they will do so in future. These observations are right censored and OLS does not distinguish between uncensored and right-censored observations; [54] (3) OLS implicitly treats time varying covariates as time-invariant. Another approach would be to model the decision whether a dyad does or does not enter a treaty as a simple logit or probit-model. However, this estimation technique does not take into account the time structure of treaty diffusion.[55]

Therefore, the primary estimation technique applied in this work is a semi-parametric Cox proportional hazard model (Cox 1972) from the family of survival models, which provide estimates for factors influencing the time until a certain event occurs. In this setting, this failure is defined as the conclusion of a tax treaty in a specific dyad. The Cox model is superior to other (parametric) event history models as in this analysis time dependency is more a nuisance while the interest lies in the relationship between the covariates and the hazard rate. Furthermore, theory does not provide a clear functional form of the baseline hazard.

The estimated Cox model is thus specified as follows:

$$h(t \mid X_{ijt} Y_{ijt}) = h_0(t) \exp(\beta' X_{ijt} + \gamma' Y_{ijt}) \tag{9}$$

where $h_0(t)$ is the baseline hazard function which also absorbs effects that do not vary across dyads and is not estimated, X_{ijt} is the set of dyad control variables outlined above, where *i* and *j* are the two dyad members. Y_{ijt} a matrix containing the spatial lag variables,

[54] A Tobit estimation that explicitly takes into account censored observations could be an alternative for this problem.

[55] Yet, some authors argue that an adequately specified logit model is a feasible alternative to the survival model (Beck et al. 1998, Carter and Signorio 2007). See the Appendix for a brief overview of this discussion.

which capture the DTT signing behaviour of other focal countries. The covariates only induce proportional shifts in the baseline hazard, but do not change its shape. The parameters β and γ are estimated via maximisation of the partial likelihood (Blossfeld et al. 2007). The date from which on a dyad starts accumulating risk (of signing a tax treaty) is taken as 1925 since in this year the first modern treaty was signed between countries that still exist nowadays, namely, Italy and Germany. However, because DTTs can only be signed between two independent tax jurisdictions, a dyad enters the analysis only in the year in which the latter of both countries became independent.[56] For such late entering countries, the amount of risk accumulated in their first year in the analysis has to be determined. The basic question is whether their risk accumulation starts at zero or is the same as all other dyads in this year. For example, Eritrea became independent in 1993, so all dyads containing Eritrea as one partner enter the study in 1993. A decision has to be made whether the risk of dyads with Eritrea in 1993 is the same as of independent dyads in 1925 (the first year of treaty conclusion) or the same as of independent dyads in 1993. The later assumption is taken since the international awareness of DTTs increased since the first treaty in 1925, and this is awareness is the same for independent and non-independent countries. Once a nation becomes independent, it finds itself in the competitive situation of the actual year and not in the situation back in 1925. As theory suggests, a newly independent country should have more pressure to sign tax treaties if DTTs are already widespread among competitor countries. This is reflected in the higher amount of risk accumulated in the year of independence. This argument is also supported by the fact that East European countries entered a multitude of DTTs with western European countries in the first years after the collapse of the Soviet Union and that China, as outlined above, signed a lot of treaties once it opened its economy.

The inclusion of a spatially lagged dependent variable introduces a potential, but probably small degree of endogeneity since dyads affect each other. Fully solving this problem would require a full spatial maximum likelihood Cox estimator, which is currently not available. To address this problem, the spatial lags are lagged by one year.[57] In addition,

---

[56] Information on independence of tax jurisdictions unfortunately is not available. Therefore, the year of political independence of a country is taken as independence date. In the cases where a DTT was signed before this date, the year of DTT signature is taken as year of independence.

[57] Even though this procedure might not fully solve the problem, it should at least mitigate endogeneity. Basically, the temporal lag should only be applied to the DTTs and not to the weighting matrix. However, due to computational intensity, the spatially weighted sum of DTTs (i.e. $\sum_{km \neq ij} w_{pq} y_{km}$ instead of y_{km}) is lagged by one year. This does not make any difference for time invariant weighting matrices such as common region.

this is reasonable since the spatial effect is unlikely to occur instantaneously, but countries need some time to react. Another potential problem is typically referred to as spatial clustering or unobserved spatial heterogeneity. This occurs since spatial patterns may not be due to spatial dependence, but can be due to the fact that close dyads are more likely to be similar than more distant dyads. In this case, the spatial lags could spuriously capture effects which have nothing to do with spatial dependence (Plümper and Neumayer 2010). Ideally, one would include dyad fixed effects to control for time-invariant dyad specific heterogeneity in the propensity to sign a tax treaty. However, this is not feasible since the Cox proportional model essentially performs a binary-outcome analysis at each point in time where at least one DTT is concluded (Cleves et al. 2004). Therefore, insufficient degrees of freedom are available to include a full set of dyad-dummies. As an alternative, as already mentioned above, two variables measuring the total number of DTTs signed by country *i* and country *j*, respectively, in year *t-1* to control for the general, but time-varying propensity of a country to enter tax treaties are included. Since this measure is time variant, it captures the overall propensity of a country towards signing DTTs more closely than a country fixed effects approach which would be feasible but would assume that this effect is constant over time. A non-constant measure is appropriate because some countries reveal principal changes in their DTT-signing behaviour with periods of inactivity followed by periods of activity and the reverse (e.g., China's first treaty with Japan in 1983 marked the beginning of a period of comprehensive treaty negotiations).

2.6. Main Results

Table 6 provides results for the Cox proportional hazard model covering the whole sample. Looking at the control variables first, the results generally are in line with expectations. Ceteris paribus, two more populous countries are always while two richer countries are in some specifications more likely to sign a DTT. Even after controlling for size (population) and the level of development (GDP per capita) of the two dyad members, there is evidence that bilateral trade increases the propensity to enter a tax treaty. Countries which already have signed a BIT are more likely to sign a DTT as well. The positive sign indicates that a BIT and a tax treaty are regarded as complements rather than substitutes. Oddly, the coefficient of RTA is negative in all specifications, but not statistically significant. A dyad,

---

The other two weighting matrices, namely export market and export product simililarity, are highly time consistent, minimising the potential bias.

in which one or both countries are classified as an offshore financial centre by the OECD, is statistically significant less likely to sign a DTT in all model specifications. If two countries have diplomatic representations, they face an increased hazard of signing a DTT, which could be either due to closer political ties in general or due to reduced transaction costs. On the other hand, more distant countries are less likely to enter a DTT. The stability of political systems, as measured by the product of the Political Constraints variable, exhibits a positive and highly significant influence on the hazard to sign a tax treaty. OECD members are generally richer, trade more and have closer political ties. After controlling for these factors and their general propensity to enter tax treaties, dyads consisting of two OECD countries are not more likely to sign a DTT, while dyads with one OECD member and a non-member are less likely to sign a DTT compared to a dyad with two non-OECD members. This is somewhat unexpected, but is driven by the covariates: once the controls for population size and GDP per capita as well as for their DTT history are removed, the dummy for an OECD-OECD dyad becomes positive and highly significant throughout. However, the OECD-non-OECD dummy variable remains negative and statistically significant. Newly independent countries have a higher propensity to sign DTTs. The higher the maximum number of DTTs signed by either of the two dyad members, the lower the likelihood of signing a DTT, consistent with theoretical expectations. A higher propensity to sign DTTs, as captured by the cumulative number of DTTs signed by either dyad member, increases the likelihood that this dyad will sign a DTT as well.

Table 6: Estimation results of Cox proportional hazard model

Model	I	II	III	IV	V	VI	VII
Spatial Lags							
W: Common region (t-1)	1.419***			1.483***	1.353***		1.421***
	(5.51)			(5.73)	(5.29)		(5.51)
W: Export market similarity (t-1)		6.143		-4.287		5.577	-4.507
		(1.07)		(-0.78)		(0.97)	(-0.81)
W: Export product similarity (t-1)			11.06***		9.965***	11.00***	9.982***
			(4.00)		(3.55)	(3.97)	(3.56)
Control Variables							
Product of populations (ln)	0.117***	0.127***	0.0945***	0.114***	0.0909***	0.0982***	0.0878***
	(4.18)	(4.50)	(3.33)	(4.01)	(3.19)	(3.41)	(3.04)
Product of GDPs per capita (ln)	0.0560*	0.0728**	0.0407	0.0538*	0.031	0.043	0.0287
	(1.96)	(2.56)	(1.39)	(1.87)	(1.05)	(1.47)	(0.97)
Bilateral trade (ln, t-1)	0.126***	0.120***	0.118***	0.125***	0.123***	0.118***	0.123***
	(7.03)	(6.88)	(6.69)	(7.05)	(6.83)	(6.67)	(6.85)
Product of openness' to trade	6.6e-05***	6.6e-05***	6.0e-05***	6.5e-05***	6.1e-05***	6.1e-05***	6.0e-05***
	(10.45)	(10.53)	(9.46)	(10.25)	(9.62)	(9.64)	(9.45)
BIT	1.322***	1.359***	1.356***	1.326***	1.313***	1.348***	1.317***
	(16.28)	(16.84)	(17.03)	(16.32)	(16.38)	(16.93)	(16.42)
RTA	-0.161	-0.132	-0.125	-0.170*	-0.144	-0.112	-0.154
	(-1.58)	(-1.31)	(-1.24)	(-1.68)	(-1.43)	(-1.13)	(-1.53)
OFC	-0.441***	-0.416***	-0.439***	-0.446***	-0.452***	-0.433***	-0.457***
	(-3.85)	(-3.68)	(-3.98)	(-3.88)	(-4.05)	(-3.92)	(-4.08)
Diplomatic representation	1.157***	1.110***	1.126***	1.155***	1.174***	1.131***	1.171***
	(11.91)	(11.27)	(11.43)	(11.88)	(12.12)	(11.51)	(12.08)
Distance (ln)	-0.273***	-0.364***	-0.379***	-0.277***	-0.280***	-0.367***	-0.284***
	(-5.56)	(-7.80)	(-8.18)	(-5.65)	(-5.66)	(-7.85)	(-5.77)
Product of political constraints	0.718***	0.752***	0.736***	0.728***	0.686***	0.719***	0.697***
	(4.57)	(4.88)	(4.82)	(4.65)	(4.43)	(4.73)	(4.52)
Dummy for OECD-OECD dyad	-0.286	0.0786	0.146	-0.294	-0.221	0.135	-0.23
	(-1.26)	(0.37)	(0.70)	(-1.30)	(-0.98)	(0.65)	(-1.03)
Dummy for OECD-non-OECD dyad	-0.609***	-0.412***	-0.341***	-0.603***	-0.558***	-0.362***	-0.551***
	(-6.31)	(-4.44)	(-3.76)	(-6.20)	(-5.78)	(-3.94)	(-5.67)
Min. years of independence	-0.0064***	-0.0078***	-0.0069***	-0.0066***	-0.0055***	-0.0066***	-0.0057***
	(-3.38)	(-4.09)	(-3.66)	(-3.45)	(-2.88)	(-3.47)	(-2.95)
Max. number of DTTs (t-1)	-0.0331***	-0.0362***	-0.0372***	-0.0334***	-0.0336***	-0.0366***	-0.0339***
	(-8.66)	(-9.47)	(-10.10)	(-8.60)	(-9.00)	(-9.77)	(-8.94)
Cumulative number of DTTs country i (t-1)	0.0418***	0.0442***	0.0438***	0.0421***	0.0411***	0.0433***	0.0414***
	(10.81)	(11.40)	(11.70)	(10.67)	(10.68)	(11.21)	(10.55)

Table 6: Estimation results of Cox proportional hazard model (continued)

Model	I	II	III	IV	V	VI	VII
Cumulative number of DTTs country j (t-1)	0.0405***	0.0431***	0.0432***	0.0407***	0.0402***	0.0427***	0.0405***
	(10.79)	(11.49)	(11.98)	(10.69)	(10.90)	(11.54)	(10.80)
Observations	212,244	212,244	212,244	212,244	212,244	212,244	212,244
DTT conclusions covered	1,385	1,385	1,385	1,385	1,385	1,385	1,385

Notes: W denotes the weighting matrix used; Coefficients displayed; Robust standard errors clustered on country dyads; Z-values in parenthesis; Breslow approximation for tied events; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

Turning to the variables capturing spatial dependence, the first three models estimate the effect of each of the spatial lags separately. Yet, the effects are not necessarily mutually exclusive and can occur simultaneously. Therefore, in models IV to VII, different combinations of the spatial effects are estimated. For treaties signed by countries in the same regions, the effect of an increase by one standard deviation (0.081) is to raise the hazard of signing a DTT by 12.2 percent in model I. For the export product similarity measure in model III, this effect is 73.8 percent. The spatial lag using export market similarity as a weighting matrix is insignificant throughout. Looking at model V, which includes both common region and the export product similarity, it can be seen that both effects become somewhat smaller, but remain statistically significant. When all three spatial lags are incorporated into model VII, the effects are not much different compared to the other model specifications.

It seems that countries are not only influenced by the behaviour of their regional peers, but also and, in substantive terms, much more strongly by the treaty signing of countries which export a similar basket of goods. No evidence, however, is found that policy-makers respond to DTT actions of countries which serve similar export markets. This is exactly what one would expect if the argument is correct that countries are not affected by what other countries do, with which they generally compete in third markets, but are affected by what other countries do with which they specifically compete for FDI. Most countries try to attract a specific type of FDI, such as investment in a particular economic sector, from which they can reap the most benefits, rather than investment whose finished goods are exported to a specific market. Furthermore, each country has a given set of natural endowments and domestic capabilities and is thereby able to attract a certain type of investment which can make productive use of these inputs. Similar export products thus indicate a similar industrial and endowment structure which in turn makes these countries close substitutes from a foreign investor's point of view. In other words, countries which export similar products strongly compete with each other for FDI, whereas countries which are located in the same macro-region also compete, if less so, and countries which merely export any range of goods and services to similar markets generally compete with each other, but do not necessarily compete with each other for scarce foreign capital.

2.7. Robustness Tests

The argument on the causes of spatial dependence relates to net capital-importing countries competing for scarce foreign capital. Some developed countries also fall into this category, but predominantly developing countries face an asymmetric capital position. Competition for scarce foreign capital should also be stronger among developing countries, which can benefit more from FDI and face tighter capital constraints. Even though the main specification controls for different types of dyads, the first robustness check constrains the analysis to dyads consisting of one OECD member and one non-member. The results are presented in Table 7. The significance levels and effect sizes of the controls are by and large the same as in the whole sample (Table 6). Regarding the spatial lags, the effects are in general similar to the whole sample: the spatial lag using the common region and the export product similarity weighting matrix are positive and significant, while no evidence for a (positive) effect for the spatial lag using the export market similarity can be established. The effect sizes of the two former spatial lags differ from the whole sample, with the spatial lag using the common region weighting matrix being slightly smaller and the coefficient on the spatial lag using the export product similarity weighting matrix being considerably larger. In model VII, an increase in the common region weighted spatial lag by one standard deviation raises the propensity of a dyad to sign a DTT by 10.8 percent, while the effect for the spatial lag weighted by export product similarity is 146.6 percent. This indicates that regional peers matter less in heterogeneous dyads in which competition for FDI is generally stronger (and asymmetry is higher), while treaty conclusions of countries with a similar export product structure matter more.

Table 7: Estimation results for OECD member and non-member country dyads

Model	I	II	III	IV	V	VI	VII
Spatial Lags							
W: Common region (t-1)	1.145***			1.293***	1.122***		1.266***
	(3.09)			(3.57)	(3.08)		(3.53)
W: Export market similarity (t-1)		-11.11		-19.52*		-9.57	-18.23*
		(-0.99)		(-1.89)		(-0.84)	(-1.73)
W: Export product similarity (t-1)			18.75***		18.39***	18.56***	18.05***
			(3.56)		(3.44)	(3.52)	(3.37)
Control Variables							
Product of populations (ln)	0.0996*	0.111**	0.0698	0.0927*	0.0549	0.067	0.0483
	(1.84)	(2.08)	(1.25)	(1.72)	(0.97)	(1.20)	(0.86)
Product of GDPs per capita (ln)	0.0639	0.0831	0.0359	0.051	0.0133	0.0309	0.00159
	(1.12)	(1.50)	(0.62)	(0.90)	(0.23)	(0.53)	(0.03)
Bilateral trade (ln, t-1)	0.169***	0.147***	0.156***	0.167***	0.177***	0.154***	0.176***
	(3.81)	(3.40)	(3.51)	(3.80)	(3.90)	(3.47)	(3.89)
Product of openness' to trade	7.1e-05***	7.2e-05***	6.6e-05***	7.1e-05***	6.5e-05***	6.6e-05***	6.4e-05***
	(7.15)	(7.18)	(6.49)	(6.94)	(6.36)	(6.39)	(6.15)
BIT	0.979***	1.034***	1.028***	1.010***	0.994***	1.045***	1.022***
	(8.65)	(9.04)	(9.12)	(8.88)	(8.88)	(9.25)	(9.11)
RTA	0.00268	0.16	0.0672	0.0659	-0.0487	0.104	0.00963
	(0.01)	(0.71)	(0.29)	(0.28)	(-0.20)	(0.46)	(0.04)
OFC	-0.555***	-0.552***	-0.601***	-0.552***	-0.597***	-0.600***	-0.592***
	(-3.87)	(-3.89)	(-4.25)	(-3.86)	(-4.20)	(-4.24)	(-4.18)
Diplomatic representation	0.941***	0.908***	0.910***	0.926***	0.936***	0.904***	0.923***
	(5.74)	(5.55)	(5.60)	(5.66)	(5.76)	(5.57)	(5.69)
Distance (ln)	-0.295***	-0.461***	-0.418***	-0.336***	-0.286***	-0.446***	-0.323***
	(-3.58)	(-5.88)	(-5.46)	(-4.06)	(-3.49)	(-5.72)	(-3.93)
Product of political constraints	1.214***	1.298***	1.268***	1.202***	1.182***	1.271***	1.175***
	(5.65)	(6.25)	(6.17)	(5.63)	(5.62)	(6.21)	(5.63)
Min. years of independence	-0.00413	-0.00487	-0.00325	-0.00436	-0.00285	-0.00333	-0.00304
	(-1.41)	(-1.63)	(-1.09)	(-1.49)	(-0.97)	(-1.11)	(-1.03)
Max. number of DTTs (t-1)	-0.0446***	-0.0494***	-0.0480***	-0.0460***	-0.0443***	-0.0488***	-0.0455***
	(-8.51)	(-9.39)	(-9.31)	(-8.74)	(-8.57)	(-9.37)	(-8.75)
Cumulative number of DTTs country i (t-1)	0.0578***	0.0620***	0.0600***	0.0588***	0.0567***	0.0607***	0.0576***
	(10.92)	(11.74)	(11.45)	(11.15)	(10.81)	(11.55)	(10.99)
Cumulative number of DTTs country j (t-1)	0.0540***	0.0582***	0.0572***	0.0548***	0.0538***	0.0578***	0.0545***
	(10.56)	(11.46)	(11.37)	(10.84)	(10.68)	(11.50)	(10.91)
Observations	55,633	55,633	55,633	55,633	55,633	55,633	55,633
DTT conclusions covered	620	620	620	620	620	620	620

Notes: W denotes the weighting matrix used; Coefficients displayed; Robust standard errors clustered on country dyads; Z-values in parenthesis; Breslow approximation for tied events; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

Estimations presented in Table 8 go one step further in not only restricting the sample to dyads consisting of one OECD and one non-OECD country, but also in presuming that DTTs are directed from the OECD to the developing country. As discussed above, power-based or 'coercive' (Simmons et al. 2006: 790) theories suggest that dominant capital-exporters such as the United States or Germany are able to control the agenda and start treaty negotiations according to their schedule and needs (Elkins et al. 2006). Furthermore, due the redistributive effect of tax treaties, capital exporters are able to increase their tax revenues at the expense of capital importers by signing a DTT. Therefore, in Table 8 DTT conclusions between an OECD member and a non-member are modelled as a directed dyad in which the process is initiated by the former and directed towards the latter. In such a setting, various forms of spatial contagion can be tested (Neumayer and Plümper 2010). Following the theoretical argument, the analysis is restricted to specific target contagion, in which the propensity of a non-OECD member to sign a DTT with a given specific OECD member depends on the weighted sum of DTTs signed by other non-OECD members with the very same OECD member. The results for the control variables are consistent with the previous analysis in an undirected dyad and also the results for the spatial lags are in line with the previous findings. In model VII, a one standard-deviation increase (0.14) in the spatial lag using the common region weighting matrix increases the probability of signing a DTT by 17.0 percent, whereas the corresponding effect for the spatial lag using the export product similarity weighting matrix is 47.6 percent.

Table 8: Estimation results for specific target contagion in a directed dyad dataset (from OECD member to non-member country)

Model	I	II	III	IV	V	VI	VII
Spatial Lags: specific target contagion							
W: Common region (t-1)	0.941***			1.089***	0.831***		1.120***
	(3.11)			(3.30)	(2.60)		(3.38)
W: Export market similarity (t-1)		0.72		-1.49		-1.832	-4.215**
		(0.53)		(-1.06)		(-1.06)	(-2.51)
W: Export product similarity (t-1)			2.302**		1.538	3.162***	3.269***
			(2.37)		(1.51)	(2.73)	(2.98)
Control Variables							
Product of populations (ln)	0.101*	0.129**	0.120**	0.0935*	0.0994*	0.113**	0.0768
	(1.86)	(2.38)	(2.22)	(1.70)	(1.83)	(2.06)	(1.38)
Product of GDPs per capita (ln)	0.0749	0.100*	0.0885	0.0683	0.0708	0.0807	0.0468
	(1.33)	(1.80)	(1.58)	(1.20)	(1.26)	(1.42)	(0.81)
Bilateral trade (ln, t-1)	0.161***	0.143***	0.147***	0.167***	0.161***	0.152***	0.178***
	(3.63)	(3.25)	(3.34)	(3.67)	(3.62)	(3.37)	(3.82)
Product of openness' to trade	7.7e-05***	7.8e-05***	8.0e-05***	7.8e-05***	7.9e-05***	7.9e-05***	7.9e-05***
	(6.95)	(7.06)	(7.21)	(6.71)	(7.05)	(7.08)	(6.73)
BIT	0.948***	0.990***	0.975***	0.955***	0.940***	0.984***	0.948***
	(8.15)	(8.56)	(8.37)	(8.24)	(8.02)	(8.47)	(8.14)
RTA	-0.107	-0.00211	-0.00467	-0.121	-0.0981	-4.53E-05	-0.126
	(-0.41)	(-0.01)	(-0.02)	(-0.46)	(-0.38)	(-0.00)	(-0.48)
OFC	-0.560***	-0.549***	-0.586***	-0.552***	-0.585***	-0.589***	-0.594***
	(-3.86)	(-3.82)	(-4.04)	(-3.79)	(-4.00)	(-4.06)	(-4.05)
Diplomatic representation	0.942***	0.928***	0.919***	0.950***	0.933***	0.921***	0.942***
	(5.81)	(5.68)	(5.62)	(5.86)	(5.74)	(5.64)	(5.81)
Distance (ln)	-0.333***	-0.427***	-0.425***	-0.329***	-0.341***	-0.435***	-0.337***
	(-4.03)	(-5.44)	(-5.38)	(-3.97)	(-4.14)	(-5.52)	(-4.09)
Product of political constraints	1.129***	1.203***	1.165***	1.134***	1.105***	1.175***	1.107***
	(5.13)	(5.66)	(5.45)	(5.16)	(5.02)	(5.51)	(5.01)
Min. years of independence	-0.00312	-0.00469	-0.00436	-0.00318	-0.00298	-0.00451	-0.00291
	(-1.04)	(-1.55)	(-1.44)	(-1.06)	(-0.99)	(-1.48)	(-0.96)
Max. number of DTTs (t-1)	-0.0526***	-0.0542***	-0.0605***	-0.0458***	-0.0589***	-0.0557***	-0.0471***
	(-9.44)	(-6.62)	(-9.05)	(-5.72)	(-8.89)	(-6.73)	(-5.78)
Cum. number of DTTs OECD member (t-1)	0.0618***	0.0640***	0.0639***	0.0608***	0.0622***	0.0632***	0.0598***
	(11.55)	(12.10)	(12.17)	(11.28)	(11.69)	(12.00)	(11.18)
Cum. number of DTTs non-OECD member (t-1)	0.0585***	0.0605***	0.0609***	0.0569***	0.0594***	0.0598***	0.0561***
	(11.24)	(11.53)	(11.89)	(10.72)	(11.48)	(11.48)	(10.67)
Observations	55,418	55,418	55,418	55,418	55,418	55,418	55,418
DTT conclusions covered	615	615	615	615	615	615	615

Notes: W denotes the weighting matrix used; Coefficients displayed; Robust standard errors clustered on country dyads; Z-values in parenthesis; Breslow approximation for tied events; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

A key assumption of a proportional hazard model is that the proportionality is maintained over time, that is, the size of the effect of a covariate is independent of the point in time at which the effect occurs. To test this assumption, a Grambsch and Therneau (1994) test is performed for each model to identify variables that potentially exhibit non-proportional effects. Exemplarily, the results of this test are shown in Table 9 for model specification VII including all three spatial lags. Only the spatial lag using the common region weighting matrix seems to be prone to having non-proportional effects. This might also be the case for the product of populations and the product of GDPs per capita, the trade measure, for the RTA and OFC dummy, for distance, for the dummy indicating a dyad consisting of two OECD members and of an OECD member and a non-member as well as for the minimum years of independence and the maximum number of DTTs. The critical variables are then interacted with the log of duration to explicitly model time-dependence.[58]

[58] Principally, any interaction with time can be chosen. As argued by Box-Steffensmeier and Zorn (2001) using the log of time is the most widely employed version. Applying $\sqrt{time}$ or time as multiplicators affects the results only marginally.

Table 9: Non-proportional hazard diagnostics (Grambsch and Therneau test)

	P	χ^2	Prob > χ^2
W: Common region (t-1)	0.05930	6.87	0.0088
W: Export market similarity (t-1)	0.00377	0.03	0.8597
W: Export product similarity (t-1)	-0.00020	0.00	0.9921
Product of populations (ln)	-0.08211	22.50	0.0000
Product of GDPs per capita (ln)	0.04078	4.40	0.0360
Bilateral trade (ln, t-1)	-0.08078	28.14	0.0000
Product of openness' to trade	-0.02332	1.01	0.3152
BIT	-0.00550	0.08	0.7708
RTA	0.12759	46.11	0.0000
OFC	-0.11656	51.59	0.0000
Diplomatic representation	0.01300	0.31	0.5789
Distance (ln)	-0.04642	5.58	0.0181
Product of political constraints	-0.01548	0.53	0.4648
Dummy fo OECD-OECD dyad	-0.11169	38.50	0.0000
Dummy for OECD-non-OECD dyad	-0.17590	52.73	0.0000
Min. years of independence	0.11847	26.81	0.0000
Max. number of DTTs (t-1)	0.04329	4.66	0.0309
Cumulative number of DTTs country i (t-1)	-0.01113	0.45	0.5036
Cumulative number of DTTs country j (t-1)	-0.00812	0.22	0.6398
Global test		369.62	0.0000

Notes: W denotes the weighting matrix used; All chi-squared statistics have 1 degree of freedom, except for Global test (19 df); Robust standard errors clustered on country dyads.

Table 10 presents the estimation results including time interactions for the full sample. As Model IV fails to converge in this specification, estimation results cannot be obtained. Since the variables are interacted with log-duration, the estimates of the direct effects can be interpreted as the effect of that covariate on the hazard of a treaty conclusion in the first year after entering the sample (where T = 1, i.e. in 1926 or in the year after the later of the dyad partners became independent). The coefficient of the non-interacted effect of the spatial lag variable with the common region weighting matrix is negative while the coefficient of its interaction with duration time is positive. This indicates that the effect is negative in early years, but becomes positive after some years. For Model I, this break-even point is reached after 53 years, which is in 1978. From this time on, a positive effect, which is increasing over time, prevails. 2,091 out of 2,325 treaties have been signed since 1978, including a large share of treaties involving developing countries, which are regularly net capital importers.

Briefly addressing the interacted control variables, the effect of population size is still positive and highly significant, but its effect is decreasing over time. As expected, the basic effect is considerably larger than in a specification without time interactions as in the latter an average effect over time is estimated. The classification of one or both countries as an offshore financial centre exhibits considerable non-proportional effects. While the main effect is positive, great in magnitude and highly significant, the interaction effect is negative and significant, indicating a decreasing effect over time. The effect becomes negative after 56 years, i.e. in 1981 (Model I). This might reflect a change in the attitude towards signing tax treaties with these countries, where the objections against entering DTTs with countries not willing to share information became greater over time.

Table 10: Estimation results of Cox proportional hazard model (with time interactions)

Model	I	II	III	IV	V	VI	VII
Spatial Lags							
W: Common region (t-1)	-23.99**				-20.50**		-23.81**
	(-2.45)				(-2.17)		(-2.51)
W: Common region (t-1) x ln(time)	6.030***				5.188**		6.004***
	(2.65)				(2.38)		(2.71)
W: Export market similarity (t-1)		-0.372				-1.918	-13.55**
		(-0.07)				(-0.34)	(-2.54)
W: Export product similarity (t-1)			10.89***		9.070***	11.10***	9.294***
			(4.16)		(3.45)	(4.28)	(3.55)
Control Variables							
Product of populations (ln)	1.093	1.293*	1.443*	fails to converge	1.123	1.354*	1.244*
	(1.45)	(1.76)	(1.95)		(1.53)	(1.85)	)1.69)
Product of populations (ln) x ln(time)	-0.246	-0.293*	-0.336*		-0.259	-0.316*	-0.291*
	(-1.38)	(-1.68)	(-1.92)		(-1.50)	(-1.82)	(-1.67)
Product of GDPs per capita (ln)	-2.387***	-2.348***	-2.174**		-2.334***	-2.292***	-2.304***
	(-2.70)	(-2.70)	(-2.46)		(-2.66)	(-2.61)	(-2.63)
Product of GDPs per capita (ln) x ln(time)	0.571***	0.563***	0.515**		0.552***	0.542***	0.543***
	(2.73)	(2.71)	(2.46)		(2.67)	(2.62)	(2.63)
Bilateral trade (ln, t-1)	2.774***	2.765***	2.822***		2.842***	2.818***	2.813***
	(4.41)	(4.41)	(4.50)		(4.52)	(4.49)	(4.49)
Bilateral trade (ln, t-1) x ln(time)	-0.625***	-0.623***	-0.637***		-0.641***	-0.636***	-0.634***
	(-4.26)	(-4.26)	(-4.35)		(-4.37)	(-4.35)	(-4.34)
Product of openness' to trade	6.13e-05***	6.14e-05***	5.64e-05***		5.75e-05***	5.63e-05***	5.57e-05***
	(10.03)	(10.09)	(8.98)		(9.17)	(8.94)	(8.75)
BIT	1.249***	1.302***	1.294***		1.247***	1.297***	1.260***
	(16.70)	(17.20)	(17.20)		(16.79)	(17.25)	(16.91)
RTA	-15.06***	-14.22***	-13.37***		-13.10***	-13.76***	-13.03***
	(-3.95)	(-3.98)	(-3.86)		(-3.79)	(-3.90)	(-3.74)
RTA x ln(time)	3.508***	3.324***	3.134***		3.058***	3.222***	3.036***
	(3.95)	(3.98)	(3.87)		(3.78)	(3.91)	(3.72)
OFC	10.48***	10.46***	10.36***		10.52***	10.42***	10.68***
	(5.13)	(5.23)	(5.23)		(5.17)	(5.24)	(5.25)
OFC x ln(time)	-2.604***	-2.597***	-2.578***		-2.615***	-2.593***	-2.657***
	(-5.28)	(-5.39)	(-5.39)		(-5.31)	(-5.41)	(-5.41)
Diplomatic representation	1.284***	1.193***	1.218***		1.300***	1.215***	1.292***
	(12.53)	(11.58)	(11.79)		(12.63)	(11.77)	(12.56)
Distance (ln)	0.0912	2.300**	2.371**		0.398	2.492**	0.497
	(0.07)	(2.00)	(2.01)		(0.30)	(2.13)	(0.39)
Distance (ln) x ln(time)	-0.0743	-0.629**	-0.646**		-0.149	-0.675**	-0.177
	(-0.24)	(-2.29)	(-2.29)		(-0.48)	(-2.42)	(-0.58)

Table 10: Estimation results of Cox proportional hazard model (continued)

Model	I	II	III	V	VI	VII
Product of political constraints	0.507***	0.565***	0.527***	0.491***	0.534***	0.526***
	(3.50)	(3.93)	(3.63)	(3.39)	(3.71)	(3.66)
Dummy fo OECD-OECD dyad	33.08***	27.43***	27.33***	31.89***	26.17***	32.72***
	(4.79)	(3.74)	(3.88)	(4.74)	(3.66)	(4.80)
Dummy fo OECD-OECD dyad x ln(time)	-8.037***	-6.596***	-6.552***	-7.729***	-6.276***	-7.930***
	(-4.70)	(-3.61)	(-3.74)	(-4.64)	(-3.53)	(-4.69)
Dummy for OECD-non-OECD dyad	16.56***	12.46***	12.76***	15.29***	11.91***	15.72***
	(5.44)	(5.13)	(5.08)	(5.00)	(4.87)	(5.11)
Dummy for OECD-non-OECD dyad x ln(time)	-4.107***	-3.064***	-3.124***	-3.792***	-2.921***	-3.888***
	(-5.70)	(-5.32)	(-5.25)	(-5.23)	(-5.03)	(-5.33)
Min. years of independence	-0.153***	-0.109*	-0.124**	-0.155***	-0.116**	-0.140**
	(-2.70)	(-1.90)	(-2.16)	(-2.75)	(-2.03)	(-2.47)
Min. years of independence x ln(time)	0.0357***	0.0249*	0.0287**	0.0362***	0.0268**	0.0327**
	(2.65)	(1.83)	(2.10)	(2.71)	(1.97)	(2.43)
Max. number of DTTs (t-1)	-0.0731	-0.0402***	-0.0403***	0.0311	-0.0404***	0.0193
	(-0.59)	(-11.25)	(-11.53)	(0.53)	(-11.41)	(0.33)
Max. number of DTTs (t-1) x ln(time)	0.00874			-0.0157		-0.0132
	(0.30)			(-1.15)		(-0.96)
Cumulative number of DTTs country i (t-1)	0.192*	0.0518***	0.0506***	0.0467***	0.0508***	0.0479***
	(1.83)	(15.26)	(15.29)	(13.76)	(14.83)	(13.78)
Cumulative number of DTTs country i (t-1) x ln(time)	-0.0339					
	(-1.38)					
Cumulative number of DTTs country j (t-1)	0.0828	0.0505***	-0.00457	0.0455***	0.0499***	0.0466***
	(0.80)	(15.24)	(-0.12)	(13.88)	(15.06)	(13.95)
Cumulative number of DTTs country j (t-1) x ln(time)	-0.0085		0.0129			
	(-0.35)		(1.36)			
Observations	212,244	212,244	212,244	212,244	212,244	212,244
DTT conclusions covered	1,385	1,385	1,385	1,385	1,385	1,385

Notes: Coefficients displayed; Robust standard errors clustered on country dyads; Z-values in parenthesis; Breslow approximation for tied events; Duration time used for time-interactions; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

2.8. Conditional Spatial Policy Dependence

In addition to the robustness tests reported above, the analysis presented in the previous section can be extended. If the argument is correct that competition for scarce foreign capital drives spatial dependence in the diffusion of DTTs, then the strength of the spatial dependence effect should systematically differ across countries in predictable ways. In other words, the strength of the spatial effects should increase or decrease depending on the values of other variables, i.e. should be conditioned by these other variables. Specifically, the number of existing DTTs, the year of a country's independence and general trade openness should condition the spatial effects. The spatial effect should be larger for countries that have signed only a few tax treaties, because countries with an extensive DTT network are more likely to have covered all important capital exporting countries already and are therefore less subject to competitive pressure. Similarly, the spatial effect should be larger in the early years of a country's independence. As long as a country is not independent, it cannot sign tax treaties;[59] however, over the course of time, other independent competitor countries can sign tax treaties and gain a competitive advantage. Once a nation becomes independent, it is 'dropped' into the actual competitive situation. It is therefore expected that the pressure to sign tax treaties is stronger right after gaining independence as a country will try to equalise its competitive disadvantage. Finally, the spatial effect should be stronger the more open a country is to trade. A positive interaction effect is anticipated as the competitive pressure should increase the more open an economy is. All else equal, a country, which is more open to foreign trade, should be more exposed to competitive pressure from abroad, should be more eager to attract foreign capital and more willing to sign international treaties.

Tables 11 to 13 test these conditional spatial policy dependence arguments, based on modified versions of models I to III (coefficients of control variables are similar and therefore not shown for reasons of space).[60] Table 11 reports results from an analysis of whether the strengths of the spatial effects depend on the number of treaties a country already has signed. The conditioning variable measures the higher of the number of treaties signed by either of the dyad members (Number of DTTs (max)). All coefficients of the

[59] As discussed above, tax treaties can be signed between independent tax jurisdictions. Since comprehensive data is not available, the year of political independence is taken as a proxy.

[60] However, as noted by Ai and Norton (2003) the straightforward interpretation of interaction effects in linear models does not extend to non-linear models such as the Cox proportional model used here. Rather, the magnitude of the interaction effects in non-linear models depends on the values of the covariates in the model and can have different signs for different observations. Therefore, in the Appendix the interaction effects are reestimated with a logit model for which appropriate tools for a detailed analysis the interaction effects are available.

spatial lag variables are positive and significant, providing the spatial effect in dyads, in which neither of the dyad member countries has previously signed any treaties. Most importantly, in line with theoretical predictions, the interaction effects are negative and significant throughout, indicating that the competitive pressure decreases as the number of treaties signed increases.

Table 11: Estimation results for spatial lags interacted with maximum number of DTTs

Model	I	II	III
W: Common region (t-1)	4.504***		
	(8.07)		
W: Common region (t-1) x max. DTT (t-1)	-0.0541***		
	(-5.60)		
W: Export market similarity (t-1)		19.25***	
		(3.19)	
W: Export market similarity (t-1) x max. DTT (t-1)		-0.210***	
		(-7.14)	
W: Export product similarity (t-1)			19.94***
			(6.64)
W: Export product similarity (t-1) x max. DTT (t-1)			-0.192***
			(-7.53)
Max. number of DTTs (t-1)	-0.0258***	-0.0126**	-0.0125**
	(-6.47)	(-2.42)	(-2.51)
Observations	212,244	212,244	212,244
DTT conclusions covered	1,385	1,385	1,385

Notes: To save space, coefficients of control variables not displayed, model specification as in Table 6; W denotes the weighting matrix used; Robust standard errors clustered on country dyads; Z-values in parenthesis; Breslow approximation for tied events; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

Table 12 reports results from testing the hypothesis that the spatial effects are conditioned by years since independence. The coefficients of spatial lags using the common region and the export product similarity weighting matrix are positive and significant, giving the spatial effects in dyads in which one of the dyad members has just gained independence. As expected, the interaction effects are negative and statistically significant at least at the ten percent level. The effect of tax treaties signed by other countries is the higher the shorter the period since independence and decreases over time which supports the argument that newly independent countries are subject to a higher pressure to enter DTT negotiations.

Table 12: Estimation results for spatial lags interacted with years since independence

Model	I	II	III
W: Common region (t-1)	2.469***		
	(5.69)		
W: Common region (t-1) x min. years of independence	-0.0242***		
	(-2.72)		
W: Export market similarity (t-1)		7.825	
		(1.36)	
W: Export market similarity (t-1) x min. years of independence		-0.0684*	
		(-1.74)	
W: Export product similarity (t-1)			13.46***
			(4.31)
W: Export product similarity (t-1) x min. years of independence			-0.0619*
			(-1.82)
Min. years of independence	-0.000792	0.000223	0.00145
	(-0.33)	(0.05)	(0.31)
Observations	212,244	212,244	212,244
DTT conclusions covered	1,385	1,385	1,385

Notes: To save space, coefficients of control variables not displayed, model specification as in Table 6; W denotes the weighting matrix used; Robust standard errors clustered on country dyads; Z-values in parenthesis; Breslow approximation for tied events; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

Table 13 finally presents results for the product of trade openness of the two countries as the conditioning variable. The coefficients of the spatial lag variables are positive and significant for common region and export product similarity, showing the effect for the (non-existing) case in which both countries are fully autarchic. The interaction effect for the common region weighted spatial lag variable is significantly positive, in line with expectations. However, unexpectedly the interaction effect is significantly negative for the export product similarity weighted spatial lag. All in all, there is thus no evidence for a higher competitive pressure for more open economies.

Table 13: Estimation results for spatial lags interacted with openness to trade

Model	I	II	III
W: Common region (t-1)	1.033***		
	(3.24)		
W: Common region (t-1) x product of openness'	5.86e-05**		
	(2.42)		
W: Export market similarity (t-1)		6.146	
		(1.05)	
W: Export market similarity (t-1) x product of openness'		-0.000000259	
		(-0.00)	
W: Export product similarity (t-1)			12.36***
			(4.37)
W: Export product similarity (t-1) x product of openness'			-0.000203**
			(-2.23)
Product of openness' to trade	5.43e-05***	6.59e-05***	0.000101***
	(6.72)	(3.21)	(5.48)
Observations	212,244	212,244	212,244
DTT conclusions covered	1,385	1,385	1,385

Notes: To save space, coefficients of control variables not displayed, model specification as in Table 6; W denotes the weighting matrix used; Robust standard errors clustered on country dyads; Z-values in parenthesis; Breslow approximation for tied events; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

2.9. Conclusion

The last decades have witnessed a wide spread of double taxation treaties with an increasing spread of diffusion. Many of these treaties involve a developing country as a signatory partner, where these countries regularly are in the role of net-capital importers. The present system of tax treaties exhibits, due to the strong reliance on the OECD model treaty, a considerable resident bias in terms of taxation rights. This means that net-capital importers can loose considerable amounts of tax revenues when entering such a treaty with dominant capital-exporting countries. Or, as stated by Irish (1974: 292), the scheme "of tax treaties creates the anomaly of aid in reverse – from poor to rich countries". This fact, however, has not tarnished the popularity of DTTs in asymmetric dyads where one country exports sizably higher amounts of capital into the partner country than the FDI that flows in the reverse direction. Several arguments have been brought forward to explain this seemingly irrational behaviour by many countries which seem, at least in terms of tax revenues, not to profit from these treaties. Most importantly, it was argued that this group of countries finds itself in a situation which can be described as a prisoners' dilemma. While it is not beneficial for developing countries as a whole to enter treaties with the standard provisions, each single country has an incentive to sign in order to gain a competitive edge in the global competition for capital. Furthermore, if other focal countries have signed tax treaties with major capital exporters, a single country has an even stronger incentive to conclude a DTT to offset its inferior position. In this chapter, the theoretical prediction that treaty conclusion by competing countries affect the DTT-behaviour of others was tested empirically. Three different concepts of competing countries were employed: (1) countries in the same region, (2) countries that export to similar markets and (2) countries that export a similar basket of goods.

In line with theoretical expectations, robust evidence is found that the treaty conclusion between two countries is in fact positively influenced by tax treaties that are negotiated between competing countries. However, a striking result of this analysis was that countries are only influenced by other countries which export a similar basket of goods, but not by those serving similar export markets. This indicates that only the former group is regarded as competitors for foreign capital as countries strive to attract a certain type of FDI rather than a broad range of investors which export finished products to a specific market.

The consequences of this behaviour can be detrimental for the group of net-capital importing countries: First, they might loose tax revenues. Second, the competitive advantage in the rivalry for FDI a net-capital importer gained by signing such a treaty is neutralised once competing countries also entered a DTT with the same major capital exporter. Collective action of developing countries could solve the dilemma: however, as argued above, this is unlikely given the large number of involved parties and due to the fact that no single country could gain a (temporary) competitive edge under a multilateral treaty. Model treaties that are more favourable for developing countries, on the other hand, are also very unlikely as experience shows: both the ANDEAN and the ASEAN model treaty are virtually non-existent on a global basis and even the UNCTAD model treaty became closer to the OECD model over time.

Such spatial dependence implies that the diffusion of DTTs becomes a quasi-automatic process as there is positive contagion: the more dyads have signed, the higher the pressure on the remaining dyads to also sign one. This process is likely to stop when net capital-importing countries believe that they will no longer lose out by refusing to sign further DTTs or, conversely, believe they can no longer benefit from signing further treaties. Difficult as a conclusive judgement on this is, a termination of the diffusion process would appear to be a long way off still.

This work provided a first insight into the global influences in the expansion of the tax treaty network. Further research should look at the level of withholding taxes agreed in the treaties as these are, together with the level of asymmetry of FDI stocks between the two contracting states, the main determinant of the potential loss in tax revenues faced by the net-capital importer. By the same token as argued above, competition for foreign capital ceteris paribus should decrease the level of withholding taxes negotiated in asymmetric dyads.

References

Ai, Chunrong and Edward C. Norton (2003), Interaction terms in logit and probit models, *Economics Letters*, 80, pp. 123-129.

Anselin, Luc (1988), *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer.

Arnold, Brian J., Jacques Sasseville and Eric Zolt (2002), Summary of the Proceedings of an Invitational Tax Seminar on Tax Treaties in the 21st Century, *Canadian Tax Journal*, 50(1), pp. 65 - 89.

Avi-Yonah, Reuven S. (1996), The Structure of International Taxation: A Proposal for Simplification, *Texas Law Review*, 75, pp. 1301-1360.

Avi-Yonah, Reuven S. (2009), Double Tax Treaties: An Introduction, in Sauvant, K.P. and L.E.Sachs (eds.), 2009, *The Effect of Treaties on Foreign Direct Investment: Bilateral Investment Treaties, Double Taxation Treaties and Investment Flows,* Oxford: Oxford Univ. Press, pp. 99-106.

Axelrod, R.M. (1984), *The Evolution of Co-operation*, New York: Basic Books.

Baistrocchi, Eduardo (2008), The Use and Interpretation of Tax Treaties in the Emerging World: Theory and Implications, *British Tax Review*, 4, 352-391.

Barthel, Fabian, Matthias Busse and Eric Neumayer (2010), The Impact of Double Taxation Treaties on Foreign Direct Investment: Evidence from Large Dyadic Panel Data, *Contemporary Economic Policy*, 28(3), pp. 366-377.

Basinger, S. J. and M Hallerberg (2004), Remodeling the Competition for Capital: How Domestic Politics Erases the Race to the Bottom, *American Political Science Review,* 98(2), pp. 261-276.

Bayer, Reşat (2006), Diplomatic Exchange Data Set, v2006.1, online: http://correlatesofwar.org.

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker (1998), Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable, *American Journal of Political Science*, 42(4), pp.1260-1288.

Bennett, D.S. and A. Stam (2005), EUGene – Expected Utility Generation and Data Management Program, available online at: http://eugenesoftware.org/.

Besen, Stanley M. and Joseph Farrell (1994), Choosing How to Compete: Strategies and Tactics in Standardization, *Journal of Economic Perspectives*, 8(2), pp. 117-131.

Blomström, Magnus (1986), Foreign Investment and Productive Efficiency: The Case of Mexico, *Journal of Industrial Economics*, 35(1), pp. 97-110.

Blonigen, B. A., and Ronald B. Davies (2004), The Effects of Bilateral Tax Treaties on U.S. FDI Activity, *International Tax and Public Finance*, 11, pp. 601-622.

Blonigen, Bruce A., and Ronald B. Davies (2005), Do Bilateral Tax Treaties Promote Foreign Direct Investment?, in: Choi, E. K., and J. C. Hartigan (eds.), 2005, *Handbook of International Trade, Volume II*, London: Blackwell, pp. 526-546.

Blonigen, Bruce A., Ronald B. Davies, Glen R. Waddell, and Helen T. Naughton (2007), FDI in space: Spatial autoregressive relationships in foreign direct investment, *European Economic Review*, 51, pp. 1303-1325.

Blossfeld, Hans-Peter, Katrin Golsch and Götz Rowher (2007), *Event History Analysis with Stata*, London: Lawrence Erlbaum Associates.

Bond, Eric W. and Larry Samuelson (1988), Strategic Behaviour And The Rules For International Taxation Of Capital*, The Economic Journal*, 99, pp. 1099-1111.

Box-Steffensmeier, Janet M. and Bradford S. Jones (2004), *Event History Modeling: A Guide for Social Scientists*, Cambridge: Cambridge University Press.

Box-Steffensmeier, Janet M. and Christopher J.W. Zorn (2001), Duration Models and

Proportional Hazards in Political Science, *American Journal of Political Science*, 45(4), pp. 972-988.

Brauner, Yarif (2003), An International Tax Regime in Crystallization, *Tax Law Review*, 56(2), pp. 259-328.

Brooks, Kim (2008), Tax Sparing: A Needed Incentive for Foreign Investment in Low-Income Countries or an Unnecessary Revenue Sacrifice?, *Queen's Law Journal*, 34(2), pp. 505-564.

Brooks, Sarah M. (2007), When Does Diffusion Matter? Explaining the Spread of Structural Pension Reforms Across Nations, *Journal of Politics*, 69(3), pp. 701-715.

Buss, Terry F. (2001), *The Effect of State Tax Incentives on Economic Growth and Firm Location Decision: An Overview of the Literature*, Economic Development Quarterly, 15(1), pp. 90-105.

Busse, Matthias and José Luis Groizard (2006), Foreign Direct Investment, Regulations, and Growth, *World Bank Policy Research Working Paper 3882*, Washington DC: The World Bank.

Carter, David B. and Curtis S. Signorio (2007), Back to the Future: Modelling Time Dependence in Binary Data, *mimeo*.

Caves, Richard (2007), *Multinational Enterprises and Economic Analysis*, Third Edition, Cambridge, UK: Cambridge University Press.

Chisik, Richard and Ronald B. Davies (2004a), Asymmetric FDI and tax-treaty bargaining: theory and evidence, *Journal of Public Economics*, 88, pp. 1119-1148.

Chisik, Richard and Ronald B. Davies (2004b), Gradualism in tax treaties with irreversible foreign direct investment, *International Economic Review*, 45(1), pp. 113-139.

Christians, Allison (2005), Tax Treaties For Investment And Aid To Sub-Saharan Africa: A Case Study, *Northwestern University School of Law Public Law and Legal Theory Papers*, 20.

Cleves, Mario A., William W. Gould and Roberto G. Guiterrez (2004), *An Introduction to Survival Analysis Using Stata*, College Station: Stata Press.

Coupé, T., I. Orlova, and A. Skiba (2008), The Effect of Tax and Investment Treaties on Bilateral FDI Flows to Transition Countries, *mimeo*.

Cox, D.R. (1972), Regression models and life-tables (with discussion), *Journal of the Royal Statistical Society, Series B*, 34, pp. 187-220.

Dagan, Tsilly (2000), The Tax Treaties Myth, *New York University Journal of International Law and Politics*, 32, pp. 939-996.

David, Paul A. (1985), Clio and the Economics of QWERTY, *American Economic Review*, 75, pp. 332-337.

Davies, Ronald B. (2003a) The OECD Model Tax Treaty: Tax Competition and Two-Way Capital Flows, I*nternational Economic Review*, 44(2), pp. 725-753.

Davies, Ronald B. (2003b), *Tax Treaties, Renegotiations, and Foreign Direct Investment*, University of Oregon Economics Working Paper No. 2003-14, Eugene: University of Oregon.

Davies, Ronald B. (2004), Tax Treaties and Foreign Direct Investment: Potential versus Performance, *International Tax and Public Finance*, 11, pp. 775-802.

Davies, Ronald B. and Thomas A. Gresik (2003), Tax Competition and Foreign Capital, *International Tax and Public Finance*, 10, pp. 127-145.

de Mooij, Ruud A. and Sjef Everdeen (2003), *Taxation and Foreign Direct Investment: A Synthesis of Empirical Research*, International Tax and Public Finance, 10, pp. 673-693.

Di Giovanni, J. (2005), What Drives Capital Flows? The Case of Cross-border M&A Activity and Financial Deepening., *Journal of International Economics*, 65, pp. 127-149.

Doernberg, Richard L. (2004), *International Taxation in a Nutshell*, St. Paul: Thomson West.

Easson, Alex (2000), Do We Still Need Tax Treaties?, *Bulletin for International Fiscal Documentation*, 54(12), pp. 619-625.

Easson, Alex (2004), *Tax Incentives for Foreign Direct Investment*, The Hague: Kluwer Law International.

Egger, P., M. Larch, M. Pfaffermayr, and H. Winner (2006), The Impact of Endogenous Tax Treaties on Foreign Direct Investment: Theory and Evidence, *Canadian Journal of Economics*, 39, pp. 901-931.

Elkins, Zachary, Andrew T. Guzman and Beth A. Simmons (2006), Competing for Capital: The Diffusion of Bilateral Investment Treaties, 1960-2000, *International Organization*, 60(4), pp. 811-846.

Findlay, Ronald (1978), Relative Backwardness, Direct Foreign Investment and the Transfer of Technology: A Simple Dynamic Model, *Quarterly Journal of Economics*, 62(1), pp. 1-16.

Finger, Joseph M. and M.E. Kreinin (1979), A Measure of 'Export Similarity' and Its Possible Uses, *The Economic Journal*, 89, pp. 905-912.

Fosfuri, Andrea, Massimo Motta and Thomas Ronde (2001), Foreign Direct Investment and Spillovers Through Workers' Mobility, *Journal of International Economics*, 53(1), pp. 205-222.

Glass, Amy Jocelyn and Kamal Saggi (2002), Multinational Firms and Technology Transfer, *Scandinavian Journal of Economics*, 104(4), pp. 495-513.

Graetz, Michael J. and Michael M. O'Hear (1997), The "Original Itent" of U.S. International Taxation, *Duke Law Journal*, 46, pp. 1021-1109.

Grambsch, Patricia M. and Terry M. Therneau (1994), Proportional Hazards Tests and Diagnostics Based on Weighted Residuals, *Biometrica*, 81, pp. 515-526.

Gordon, Roger and Wei Li (2009), Tax structures in developing countries: Many puzzles and a possible explanation, *Journal of Public Economics*, 93, pp. 855-866.

Guzman, Andrew T. (1998), Why LDCs Sign Treaties That Hurt Them: Explaining the Popularity of Bilateral Investment Treaties, *Virginia Journal of International Law*, 38, pp. 639-688.

Hartman, David G. (1985), Tax Policy and Foreign Direct Investment, *Journal of Public Economics*, 26, pp. 107-121.

Hamada, Koichi (1966), Strategic aspects of taxation on foreign investment income, *Quarterly Journal of Economics*, 80, pp. 361-375.

Hayashi, M. and R. Boadway (2001), An empirical analysis of intergovernmental tax interaction: the case of business income taxes in Canada, *Canadian Journal of Economics*, 34(2), pp. 481-503.

Hays, Jude C (2003), Globalization and Capital Taxation in Consensus and Majoritarian Democracies, *World Politics,* 56, pp. 79-113.

Hays, Jude C. (2009), *Globalization, Domestic Institutions, and the New Politics of Embedded Liberalism*, Oxford: Oxford University Press (forthcoming).

Henisz, W. J. (2000), The Institutional Environment for Economic Growth, *Economics and Politics*, 12(1), pp. 1-31.

Helpman, Elhanan (2006), Trade, FDI, and the Organization of the Firms, *Journal of Economic Literature*, 44(3), pp. 589-630.

Hochgatterer, Claudia and Markus Leibrecht (2009), Tax competition as a cause of falling corporate income taxes: A literature survey, Vienna: *Discussion Papers SFB International Tax Coordination*, 32.

International Bureau of Fiscal Documentation (2009), *Tax Treaty Database*, Online database.

Irish, Charles R. (1974), International Double Taxation Agreements and Income Taxation at Source, *International and Comparative Law Quarterly*, 23(2), pp. 292-316.

Jahn, Detlev (2006), Globalization as ‘Galton's Problem’: The Missing Link in the Analysis of Diffusion Patterns in Welfare State Development, *International Organization*, 60 (2), pp. 401-431.

Janeba, Eckhard (1995), Corporate income tax competition, double taxation treaties, and foreign direct investment, *Journal of Public Economics*, 56, pp. 311-325.

Katz, Michael L. and Carl Shapiro Carl (1985), Network Externalities, Competition and Compatibility, *American Economic Review*, 75(3), pp. 424-440.

Kaplan, E.L. and P. Meier (1958), Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, 53, pp. 457-481.

Kolm, Evelyn (2005), The Emerging System of Double Taxation Agreements in the late 19th and 20th Centuries, *Development and Finance*, pp. 66-75.

Kosters, Bart (2004), The United Nations Model Tax Convention and its Recent Developments, *Asia Pacific Tax Bulletin*, 10, pp. 4-11.

Manger, Mark S. (2006), The Political Economy of Discrimination: Modelling the Spread of Preferential Trade Agreements, Working Paper. Montreal: Department of Political Science, McGill University.

Mansfield, Edwin and Anthony Romeo (1980), Technology Transfer to Overseas Subsidiary by US-based Firms, *Quarterly Journal of Economics*, 95(4), pp. 737-750.

Markusen, James (2002), *Multinational Firms and the Theory of International Trade*, Cambridge, MA: MIT Press.

Markusen, James and Anthony Venables (1999), Foreign Direct Investment as a Catalyst for Industrial Development, *European Economic Review*, 43(2), pp. 335-356.

Musgrave, Peggy B. (1969), United States Taxation of Foreign Investment Income: Issues and Arguments, *International Tax Program*, Cambridge: Harvard Law School.

Navaretti, Giorgio and Anthony Venables (2004), *Multinational Firms in the World Economy*, Princeton, NJ: Princeton University Press.

Neumayer, Eric (2007), Do double taxation treaties increase foreign direct investment to developing countries?, *Journal of Development Studies*, 43(8), pp. 1501-1519.

Neumayer, Eric and Thomas Plümper (2009), Making Spatial Analysis Operational: Ado-files for Generating Spatial Effect Variables in Monadic and Dyadic Data, London: London School of Economics.

Neumayer, Eric and Thomas Plümper (2010), Spatial Effects in Dyadic Data, *International Organization*, 64(1), pp. 145-166.

Norton, Edward C., Hua Wang, Chunrong Ai (2004), Computing interaction effects and standard errors in logit and probit models, *The Stata Journal*, 4(2), pp. 154-167.

Ring, Diane (2007), International Tax Relations: Theory and Implications, *Tax Law Review*, 60(20), pp. 83-154.

Rodríguez-Clare, Andrés (1996), Multinationals, Linkages, and Economic Development, *American Economic Review*, 86(4), pp. 851-873.

Rork, J.C. (2003), Coveting thy Neighbours' Taxation, *National Tax Journal*, 56, pp. 775-787.

Rose, Andrew (2009), Data on bilateral trade are available at: http://haas.berkeley.edu/~arose.

Seligman, Edwin R.A. (1928), *Double Taxation and International Fiscal Cooperation,* New York: The Macmillan Company.

Shelton, Ned (2004), *Interpretation and Application of Tax Treaties*, London: LexisNexis UK.

Simmons, Beth and Zachary Elkins (2004), The Globalization of Liberalization: Policy Diffusion in the International Political Economy, *American Political Science Review*, 98(1), pp. 171-189.

Simmons, Beth, Frank Dobhin, and Geoffrey Garrett (2006), Introduction: The international Diffusion of Liberalism, *International Organization*, 60(4), pp. 781-810.

Swank, Duane (2006), Tax Policy in an Era of Internationalization: Explaining the Spread of Neoliberalism, *International Organization*, 60(4), pp. 847-882.

Thuronyi, Victor (2003), Recent Treaty Practice on Tax Sparing, *Tax Notes International*, 29, pp. 301-303.

Thuronyi, Victor (2010), Tax Treaties and Developing Countries, in: Michael Lang, Pasquale Pistone, Josef Schuch, Claus Staringer, Alfred Stock and Martin Zagler (eds.): *Tax Treaties: Building Bridges between Law and Economics*, Amsterdam: IBFD 2010, pp. 441-456.

Transparency International (2009), *Corruption Perception Index*, Online Database.

Organization for Economic Cooperation and Development (OECD) (1998), *Harmful Tax Competition: A Global Issue*, Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2004) *The OECD's Project on Harmful Tax Practices: The 2004 Progress Report*, Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2005), *Model Tax Convention on Income and Capital: Condensed Version*, Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2009), *A Progress Report on the Jurisdictions surveyed by the OECD Global Forum in Implementing the Internationally Agreed Tax Standard*, http://www.oecd.org/dataoecd/50/0/43606256.pdf, accessed: 01.12.2009.

Plümper, Thomas, Vera E. Troeger and Hannes Winner (2009), Why is there no Race to the Bottom in Capital Taxation? Tax Competition between Countries of Unequal Size, Different Levels of Budget Rigidities and Heterogeneous Fairness Norms, *International Studies Quarterly,* 53(3), pp. 761-786.

Plümper, Thomas and Eric Neumayer (2010), Model Specification in the Analysis of Spatial Dependence, *European Journal of Political Research*, 49(3), pp. 418-442.

United Nations Conference on Trade and Development (UNCTAD) (2000), *Tax Incentives and Foreign Direct Investment: A Global Survey*, ASIT Advisory Studies 16, United Nations, New York and Geneva.

United Nations Conference on Trade and Development (UNCTAD) (2007), *Bilateral Investment Treaties*, Online database.

US Senate Committee on Foreign Relations (2006), *Testimony of Patricia A. Brown, Deputy International Tax Counsel (Treaty Affairs), United States Department of the Treasury Before the Senate Committee on Foreign Relations on Pending Income Tax Agreements*, http://foreign.senate.gov/imo/media/doc/BrownTestimony060202.pdf; accessed: 18.02.2011.

Wijnen, W.F.G. (1998), Towards a new UN Model, *Bulletin for International Fiscal Documentation*, 52(3), pp. 135-143.

World Bank (2009a), *World Development Indicators*, Online Database.

World Bank (2009b), *Ease of Doing Business*, Online Database.

World Trade Organization (WTO) (2009), *Regional Trade Agreements Information System*, Online Database.

Appendix

Interaction effects using Logit estimation

As noted by Ai and Norton (2003) the interpretation of interaction effects in non-linear models is not straightforward interpretation as the effect size depends on the values of covariates in the model and can have different signs for different observations. Therefore, the interaction effects are re-estimated with a logit model, for which the marginal interaction effects are computed using the Stata command *inteff.ado* (Norton et al. 2004).

To mirror the Cox proportional hazard model, Beck et al. (1998) suggest the use of a corrected model of cross sectional data with a binary dependent variable (BCTSCS), which is estimated with a logit model. The authors propose estimating a complementary log-log regression and enriching the model specification with a full set of time dummies. However, as argued by Carter and Signorio (2007), this approach suffers from two drawbacks: First, under specific circumstances, a time dummy is a perfect predictor of Y = 1 and cannot be estimated which leads to a loss of observations. Second, especially in large T datasets, a full set of time dummies uses considerable degrees of freedom and reduces efficiency. As a remedy, they propose to include time, $time^2$ and $time^3$ as regressors.[61] More specifically, the logit model reads as follows:

$$\Pr(y_{it} = 1) = \frac{1}{1 + \exp(-(X\beta + \gamma_1 t_i + \gamma_2 t_i^2 + \gamma_3 t_i^3))} \qquad (10)$$

where $\gamma_1 t_i + \gamma_2 t^2{}_i + \gamma_3 t^3{}_i$ is a cubic approximation to the hazard which can accommodate a very large variety of shapes, while X contains both a constant as well as the spatial lags, their interactions and the control variables. Since the focus is on the parameters influencing the first signature of a tax treaty (rather than the effects on maintaining a DTT), only the first year of treaty conclusion is kept in the dataset while all subsequent years are dropped.

Table 14 shows the results for the interaction of the spatial lags with the maximum number of DTTs, where the last row displays the mean of the marginal interaction effects for all observations. As in the Cox model, the non-interacted basic effect of the maximum number of DTTs is negative and significant; the basic effects of the spatial lags are positive and

[61] Conducting a Monte Carlo Simulation, Carter and Signorio (2007) show that their approach is superior to the time dummy solution by Beck et al. (1998).

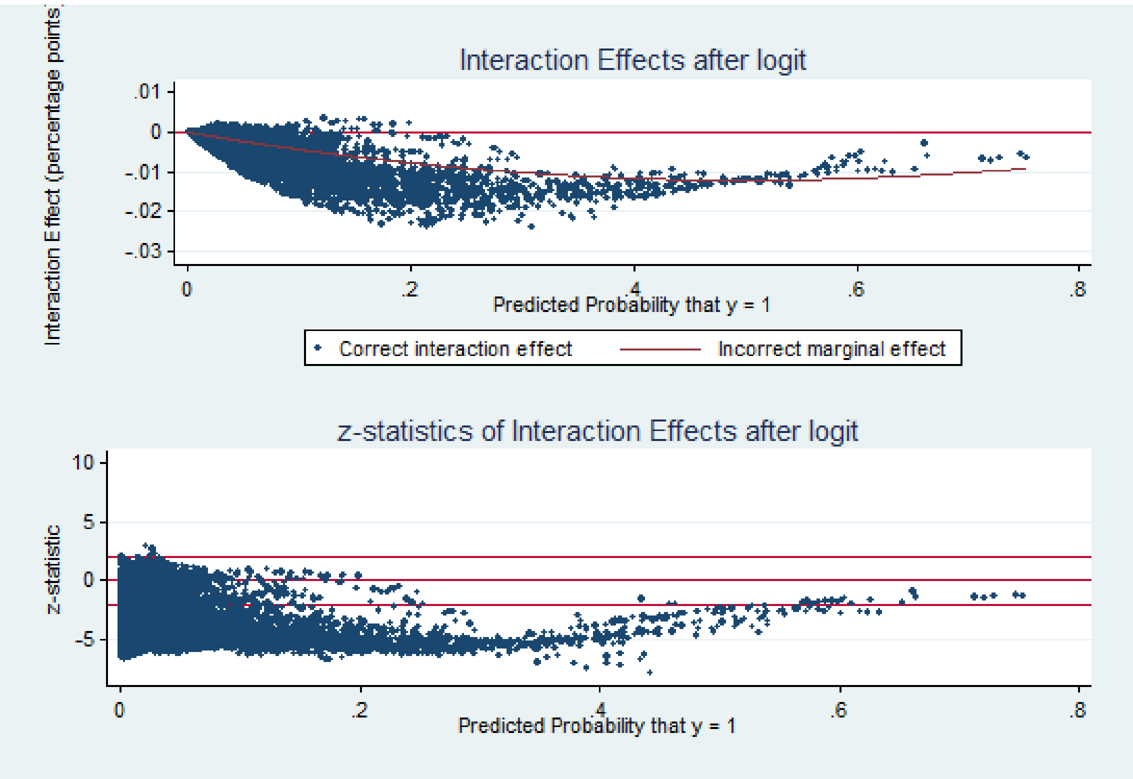
highly significant. All three interaction effects are negative and significant, confirming the previous results. Figure 14 plots the marginal interaction effect and the z-statistics against the predicted probability that a dyad concludes a DTT. The marginal effect for the interaction between the spatial lag using common region as a weighting matrix and the maximum number of DTTs is negative for virtually all observations, only for a few with a predicted probability of less than 0.2 it is positive. The significance if displayed at the bottom of the panel: Again, for most of the observations the absolute z-value is greater than the critical value, indicating statistical significance. For some dyads with a predicted probability below 0.2, however, the effect is not statistically significant at the five percent level.

Table 14: Logit estimation results: Spatial lags interacted with number of DTTs (max.)

Model	I	II	III
W: Common region (t-1)	3.708***		
	(5.43)		
W: Common region (t-1) x max. DTT (t-1)	-0.0483***		
	(-3.97)		
W: Export market similarity (t-1)		18.10***	
		(2.85)	
W: Export market similarity (t-1) x max. DTT (t-1)		-0.203***	
		(-5.86)	
W: Export product similarity (t-1)			18.27***
			(5.46)
W: Export product similarity (t-1) x max. DTT (t-1)			-0.188***
			(-6.18)
Max. number of DTTs (t-1)	-0.0377***	-0.0237***	-0.0224***
	(-7.22)	(-3.67)	(-3.54)
Mean marginal interaction effect	-.0006163	-.0030928	-.0031921

Notes: To save space, coefficients of control variables not displayed, model specification as in Table 6; W denotes the weighting matrix used ; Coefficients displayed; Robust standard errors clustered on country dyads; Z-values in parenthesis; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

Figure 14: Interaction effect between the spatial lag using common region as a weighting matrix and number of DTTs (max)



Figures 15 and 16 replicate the information for the spatial lag using export market similarity and export product similarity, respectively, as a weighting matrix. For both interactions, the effect is negative for the vast majority of observations and the overall pattern looks very similar: the effect is not statistically different from zero for quite a few observations with a predicted probability of less than 0.25 and more than 0.6.

Figure 15: Interaction effect between the spatial lag using export market similarity as a weighting matrix and number of DTTs (max)

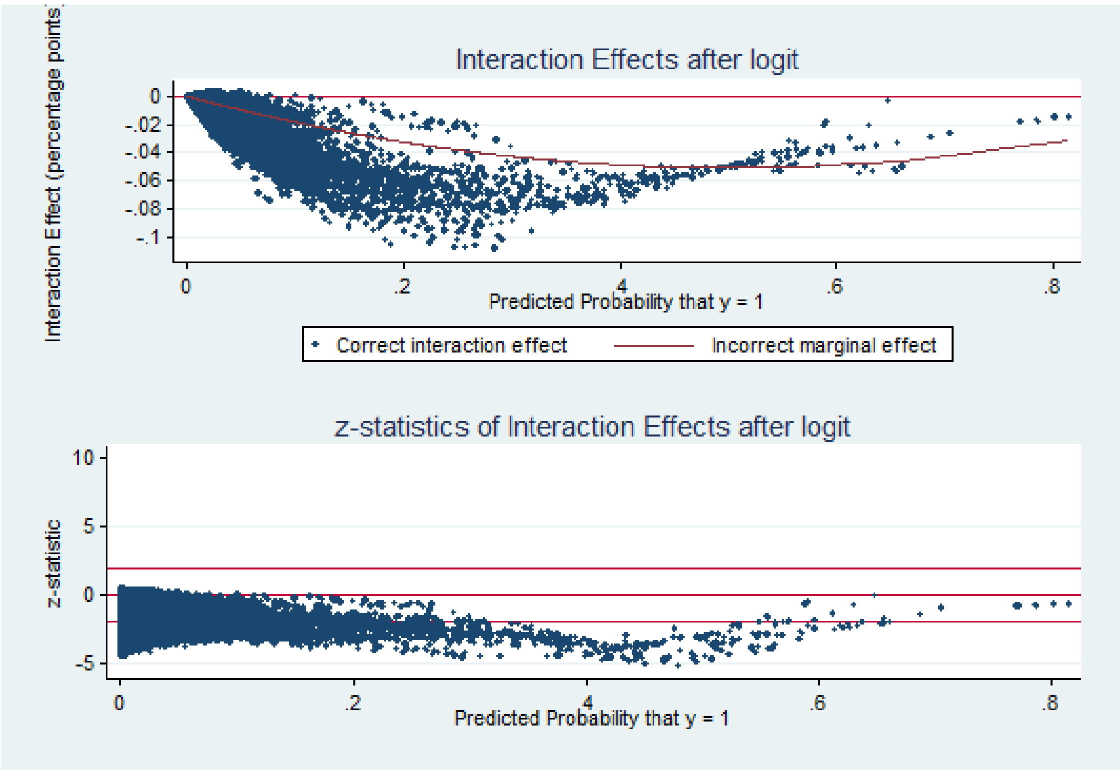


Figure 16: Interaction effect between the spatial lag using export product similarity as a weighting matrix and number of DTTs (max)

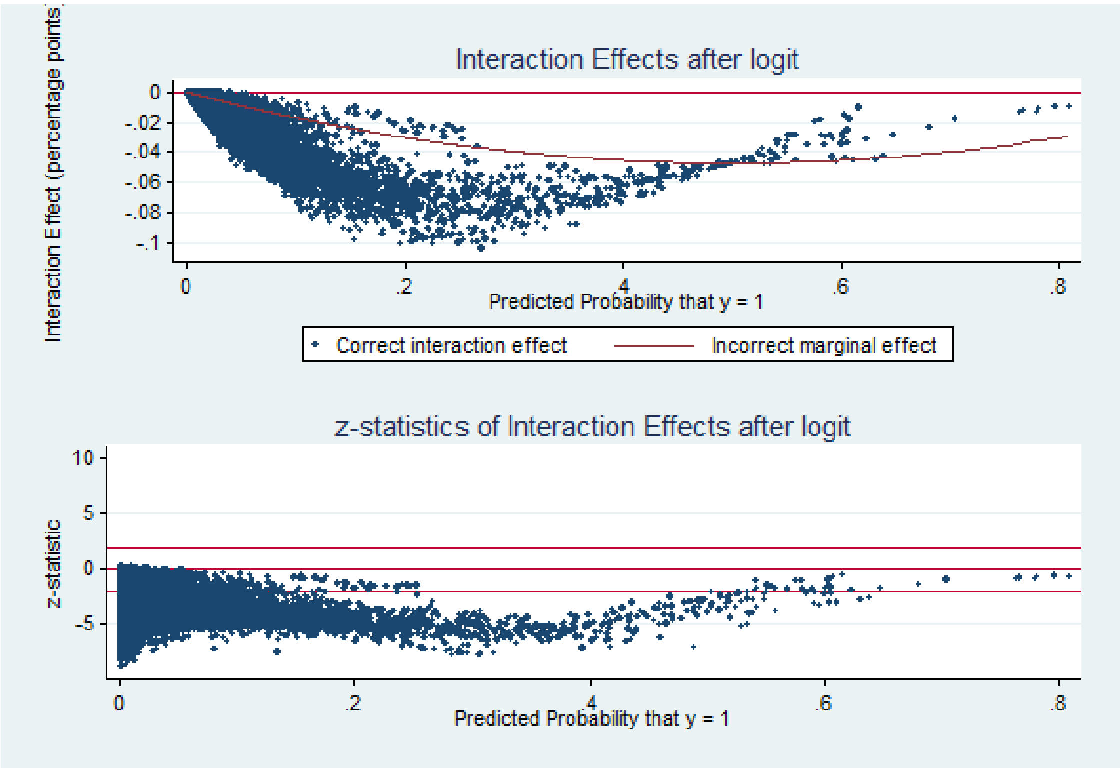


Table 15 shows the Logit estimation results for the interactions with the years since independence, while Figure 17 to 19 display the interaction effects in more detail. In line with the Cox proportional hazard model, the main effects of the spatial lags with the common region and the export product similarity weighting matrix are positive and significant, whereas all three interaction effects are negative and significant. The graphs reveal that the interaction effects are negative and significant for the vast majority of observations.

Table 15: Logit estimation results: Spatial lags interacted with years since independence

Model	I	II	III
W: Common region (t-1)	2.309***		
	(4.46)		
W: Common region (t-1) x min. years of independence	-0.0307***		
	(-2.74)		
W: Export market similarity (t-1)		8.534	
		(1.39)	
W: Export market similarity (t-1) x min. years of independence		-0.0904**	
		(-1.99)	
W: Export product similarity (t-1)			12.69***
			(3.72)
W: Export product similarity (t-1) x min. years of independence			-0.0796**
			(-2.00)
Min. years of independence	-0.00204	0.00123	0.00228
	(-0.77)	(0.24)	(0.44)
Mean marginal interaction effect	-.0002118	-.0007064	-.0007789

Notes: To save space, coefficients of control variables not displayed, model specification as in Table 6; W denotes the weighting matrix used ; Coefficients displayed; Robust standard errors clustered on country dyads; Z-values in parenthesis; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

Figure 17: Interaction effect between the spatial lag using common region as a weighting matrix and years of independence

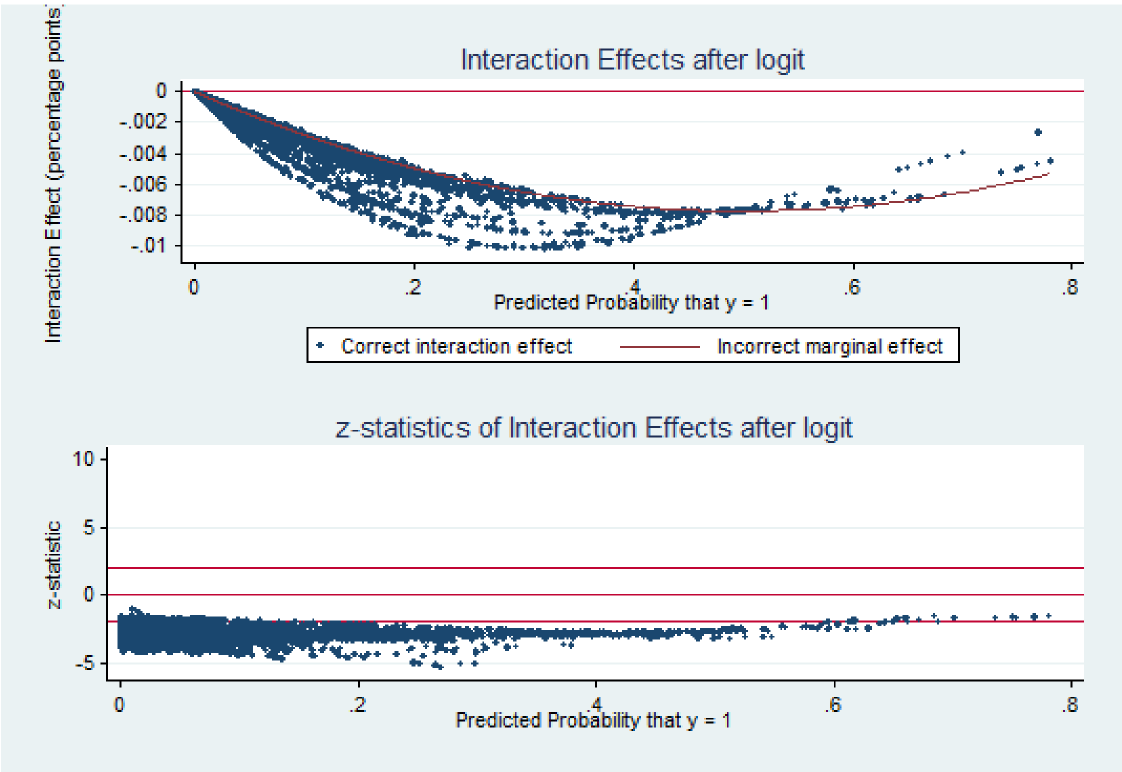


Figure 18: Interaction effect between the spatial lag using export market similarity as a weighting matrix and years of independence

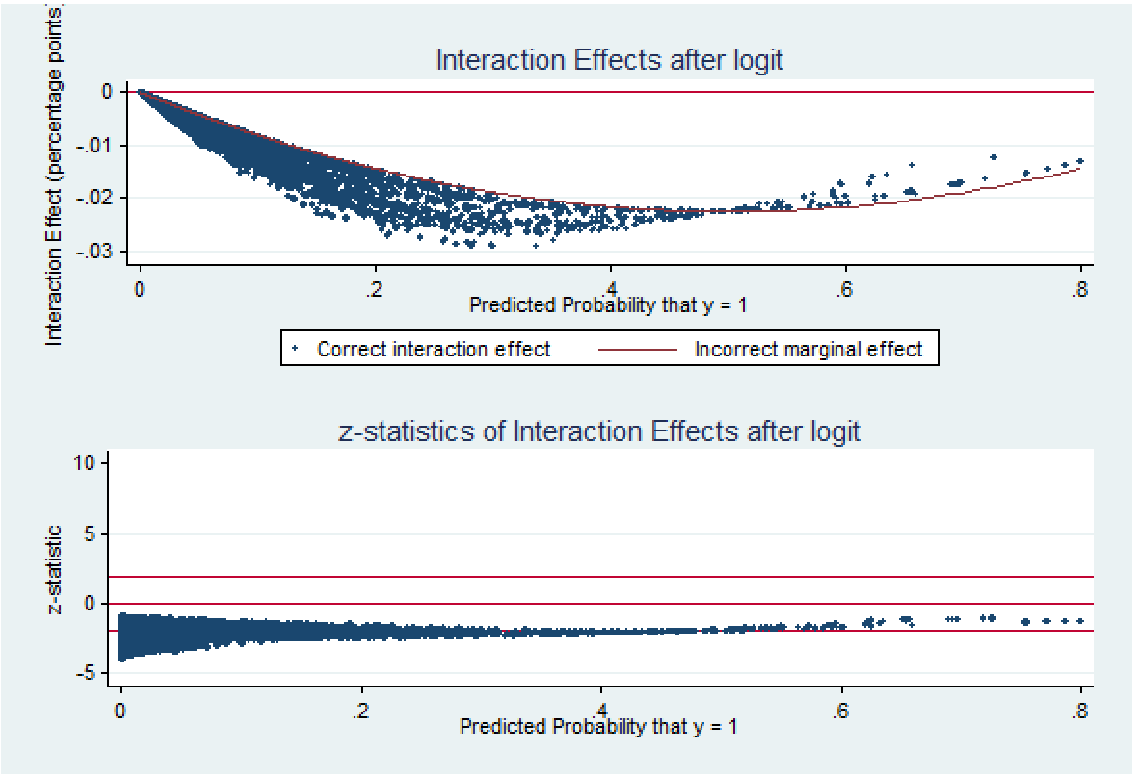
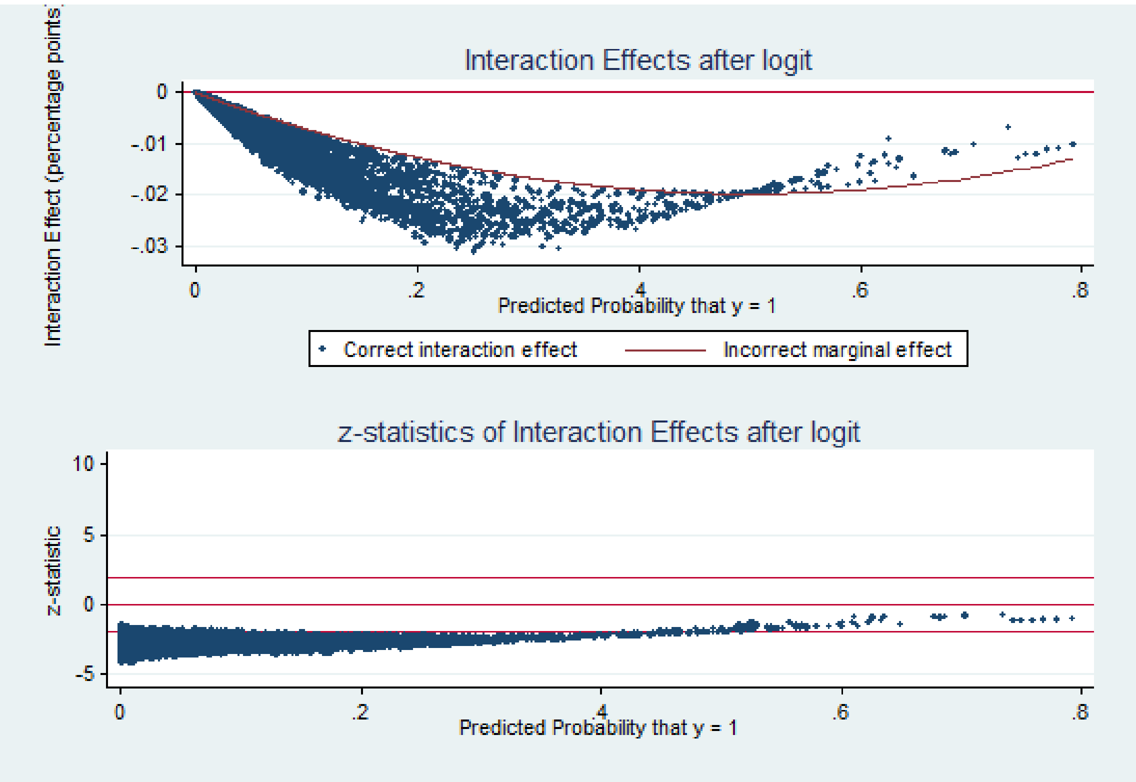


Figure 19: Interaction effect between the spatial lag using export product similarity as a weighting matrix and years of independence



Finally, Table 16 presents the interaction effect with openness to trade. As in the Cox proportional hazard model, the main effect of the product of openness' is positive and significant. Contrary to the Cox model, the spatial lag using the common region weighting matrix is not significant. The spatial lag using the export market similarity weighing matrix is negative, yet not significant, whereas the spatial lag with the export product similarity weighting matrix is positive and significant as in the Cox model. The estimated interaction effects are positive and significant throughout. Figures 20 to 22 again present the detailed analyses of the interaction effects. For virtually all observations of the spatial lag using the common region weighting matrix, the effect is positive and significant for most of the observations. Regarding the interaction effect for the spatial lag with the export market similarity, it can be seen that the effect is not significant for any observation. Finally, for the spatial lag with the export product similarity, the interaction effect is positive for the vast majority of observation which have a predicted probability of less than 0.4. However, with the exception of a few observations with a predicted probability of less than 0.2, the effect is not statistically significant at the five percent level.

Table 16: Logit estimation results: Spatial lags interacted with the product of openness to trade

Model	I	II	III
W: Common region (t-1)	0.234		
	(0.43)		
W: Common region (t-1) x product of openness'	0.0131*		
	(1.89)		
W: Export market similarity (t-1)		-2.244	
		(-0.35)	
W: Export market similarity (t-1) x product of openness'		0.0889***	
		(4.34)	
W: Export product similarity (t-1)			5.758*
			(1.74)
W: Export product similarity (t-1) x product of openness'			0.0630***
			(3.55)
Product of openness' to trade	6.90e-05***	4.05e-05***	4.16e-05***
	(7.49)	(3.13)	(3.22)
Mean marginal interaction effect	6.48e-07	2.49e-06	3.06e-06

Notes: To save space, coefficients of control variables not displayed, model specification as in Table 6; W denotes the weighting matrix used ; Coefficients displayed; Robust standard errors clustered on country dyads; Z-values in parenthesis; * statistically significant at 0.1, ** 0.05, or *** 0.01 level.

Figure 20: Interaction effect between the spatial lag using common region as a weighting matrix and product of openness'

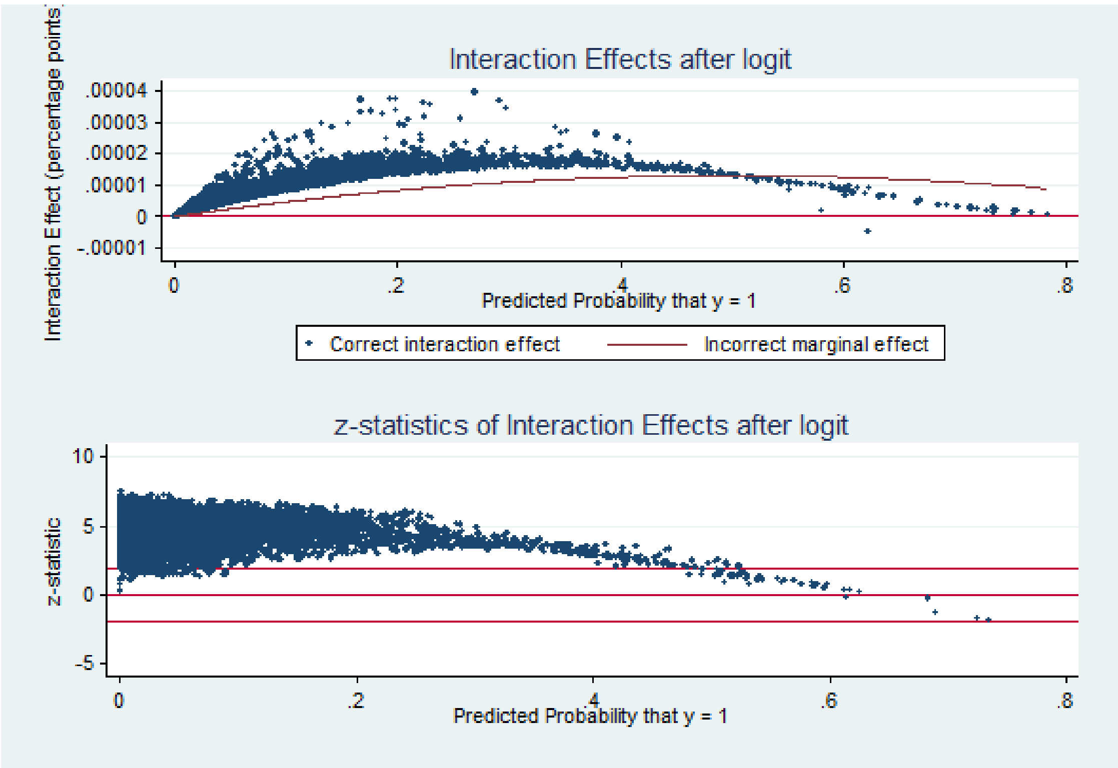


Figure 21: Interaction effect between the spatial lag using export market similarity as a weighting matrix and product of openness'

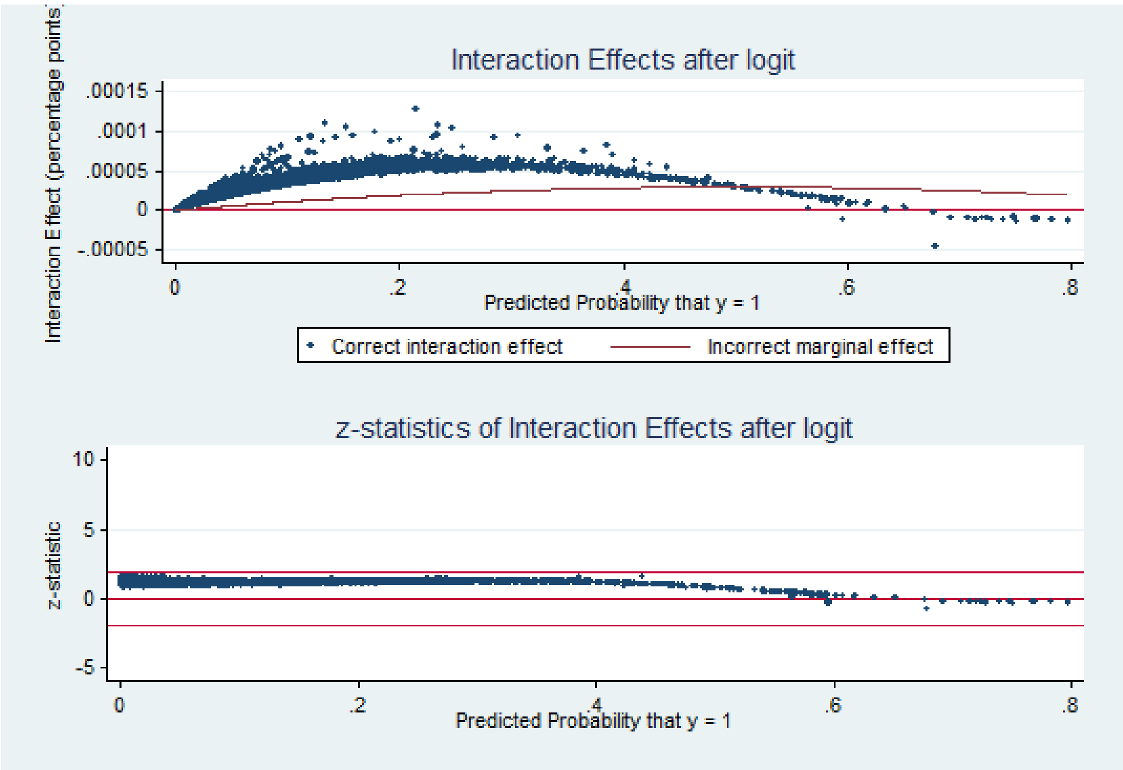
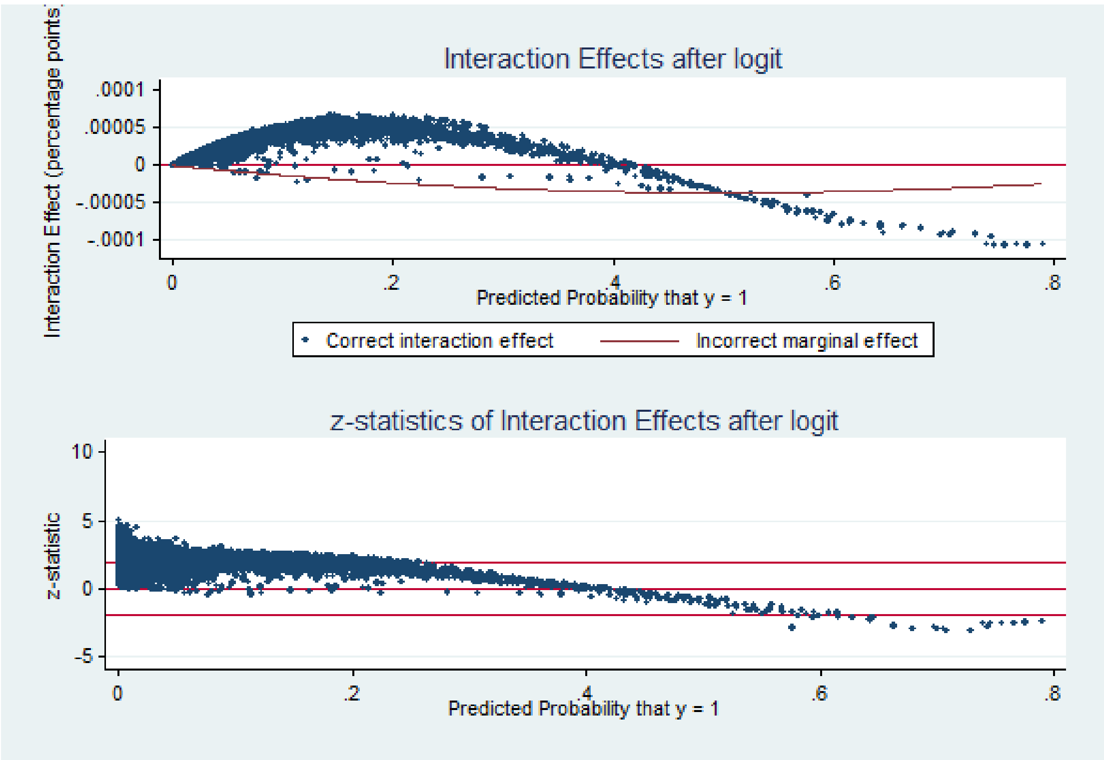


Figure 22: Interaction effect between the spatial lag using export product similarity as a weighting matrix and product of openness’



3. Are Donors Sheep? – Spatial Dependence in Bi- and Multilateral Aid Giving Patterns

3.1. Introduction

Does aid allocation depend on the aid allocation decisions of other donors? So far, this question has been answered only unsatisfactorily in the literature. Not only would an answer to this question complete our understanding of the determinants of aid giving and of international power relations; it would also enable us to assess the degree to which aid is used for achieving strategic goals. There are theoretical reasons for both positive and negative spatial dependence. Negative spatial dependence means that a given donor reduces its aid effort in a recipient country if other donors allocate more funds to this beneficiary. If such a pattern is found, it could indicate a donor specialisation over time, where each donor concentrates its activities on a set of partner countries in development cooperation. In contrast, positive spatial occurs if changes in aid allocation by various donors are in the same direction, i.e. a given donor allocates more aid to a recipient if other donors do so. Potential causes for such behaviour could be herding or group solidarity, reduction of uncertainty about the effectiveness of aid projects or strategic interaction between donors. The latter argument applies because aid is at least partly given for strategic reasons either by establishing commitment on part of the donor or through dependency on part of the recipient (McKinlay and Little 1977). Furthermore, an important donor is able to influence the behaviour of a beneficiary in its own interest. If foreign aid is used to pursue national interests and to secure the donor's sphere of influence, a donor also has to observe aid allocation decisions by other donors and take changes in their aid giving into account when allocating its own aid. National interest can be widely defined in terms of economic, military or political interest. Economic interest encompasses access to raw materials or market access, military interest include counter-terrorism and securing the loyalty of allies in conflicts, while an example for political interest is gaining a beneficiary's support in international bodies such as the UN general assembly.

If spatial dependence between donors is found, it would complement the two traditional rationales to provide ODA, first distinguished in the often cited work by McKinlay and Little (1977): donor interest and recipient need.[62] This distinction captures the idea that donors neither behave entirely selfish nor completely altruistic when allocating aid, i.e. that they pursue their national interests while also taking the need for aid in the beneficiary country into account. The humanitarian view identifies economic assistance as the primary reason for aid whereas in the foreign policy view aid is regarded as a mean to satisfy the

[62] See, for instance Maizels and Nissanke (1984), Berthélemy and Tichit (2004) and Berthélemy (2006) for empirical studies on the relative importance of these two factors for various donors.

interest of the donor (McKinlay and Little 1977). Recipient need, on the one hand, means that more aid should go to countries with the greatest need for poverty alleviation. Furthermore, natural disasters and political or ethnic conflicts can trigger need for (additional) foreign assistance. On the other hand, donor interests encompass a wide range of geopolitical, economic, military-strategic, or cultural interest.

This study examines empirically whether such spatial dependence exists in bi- and multilateral aid giving. Even though there are a few studies that include aid by other donors as an explanatory variable, these are implicitly based on the oversimplifying assumption that each donor reacts to the aid allocation of every other donor in the same way, i.e. that a given donor is influenced by any other donor identically. For instance, there is no reason to believe that bilateral donors are as much influenced by multilateral organisations as by other bilateral donors. In addition, no previous study has so far addressed the issue of strategic interaction among donors.

The contribution of this analysis is threefold: First, exploiting the previously unpublished PLAID dataset (Nielson et al. 2010) and covering 26 country donors as well as multilateral institutions, 139 recipient countries and 35 years, the empirical work is based on a larger database than any previous study. Second, the assumption of identical influence of all donors is abandoned. Third, the interaction of strategic motives by different donors is modelled explicitly. The results show that there is evidence for positive spatial contagion, that is, a donor is influenced by other donors in its aid allocation decisions. ODA transfers by different donors consequently act as complements rather than as substitutes. It is shown that donors are in fact not influenced in the same way by every other donor, but that rather the largest players in international development aid giving act as leaders whereas smaller donors take the role of followers. Little evidence is found for strategic interaction concerning economic and political interest, whereas tentative evidence suggest that military alliances of different donors with the same recipient matter. Generally, smaller donors depend more strongly on others than larger donors.

The remainder of this paper is structured as follows: in Section 2, potential causes of spatial dependence and its consequences are discussed, while Section 3 briefly reviews the relevant literature. Section 4 describes the estimation methodology, the construction of the spatial lags and the dataset used for estimation. Section 5 provides the main results both on an aggregate level as well as for individual donors while the analysis is extended to conditional spatial dependence in Section 6. Section 7 concludes.

3.2. Origin and consequences of spatial dependence

Causes of spatial dependence

There are theoretical reasons both for positive and negative spatial dependence. In the former case, a given donor will increase (decrease) its aid share to a given recipient if other donors give more (less) aid to the same recipient. Negative spatial dependence is found if a donor reduces (raises) the aid budget to a recipient country if other donors allocate more (less) aid to this recipient. Aid allocation by different donors act as complements with positive spatial dependence, whereas they are substitutes with negative spatial dependence.

The first source of positive spatial dependence is the potentially great uncertainty involved in aid projects. Even though donors might attach conditions to their aid and support, they do not have full information on how the money will be used in the recipient country. Uncertainty occurs at different levels: (1) it is not always guaranteed that the funds will actually reach the beneficiary development organisation or project, (2) the donor cannot fully influence how the money is spent by the implementing organisation and (3) there is uncertainty whether the envisaged positive development effects will materialise. A certain degree of this uncertainty might be reduced by intensive preparatory work and close monitoring of the projects; however donors often lack the administrative capacity to scrutinise the use of granted funds. At the same time, donors are accountable to the national taxpayers for spending their money responsibly. This is why some donors might follow the example of other donors and give funds to the same countries as this might be a signal for good aid projects (Vázquez 2008).[63] Doing the same as others reduces search costs for profitable aid projects on the one hand and it alleviates the risk of allegations by the taxpayer on the other hand. This undifferentiated movement by various donors in the same direction has been labelled "herding" (Frot and Santiso 2011) or "bandwagon effect" (Vázquez 2008) by other authors.

The second cause of positive spatial dependence is strategic interaction among donors. As argued by McKinlay and Little (1977), aid can be used to establish commitment on part of

[63] Reinhardt (2006) provides a micro-level study that supports this argument.

the donor and dependency on the part of the recipient country which can be used to realise different policy utilities. Commitment could either reduce the risk of intervention of a hostile state, incentivise a recipient to stay within the sphere of influence of the donor, or be used to discourage a particular country to move into a rival's sphere of influence (McKinlay and Little 1977). In addition, dependency increases the potential of the donor to control the recipient and influence its behaviour in its own interest.[64] This is true not only for political and military, but also for economic interests. If a country seeks to protect its own national interests through foreign policy, it also has to observe the actions of other countries and their ambitions to secure influence closely. Since ODA is one mean to exert and secure power, donors should therefore monitor other donors' aid allocation decisions and react correspondingly. As Frot and Sanitiso (2011: 65) put it, donors that do not "participate in the aid splurge" may fear to be left out and miss investment and diplomatic opportunities in the future.

Finally, negative spatial dependence can be a sign for increasing donor specialisation over time: Starting from a situation with a very wide-spread diffusion of ODA donors begin to concentrate their aid efforts on a certain set of recipient countries. As a consequence, these countries on the recipient-shortlist will receive a higher share of the total aid budget of a donor, while the aid share of non-partner countries will decrease over time. If this behaviour can be observed for a large number of donors – and their set of recipient countries is not identical – this leads to negative spatial dependence. Such a donor specialisation would be in line with the long history of political pledges to improve donor coordination, such as the Monterrey Consensus in 2002, the Rome Declaration on Harmonisation in 2003, the Paris Declaration on Aid Effectiveness in 2005 and the Accra Agenda for Action in 2008.[65] These political agreements were negotiated against the background of the plethora of negative effects of excessive donor fragmentation and project proliferation as well as of problems associated with a lack of coordination.[66] At the EU level, the Code of Conduct on Complementarity and Division of Labour in Development Policy stipulates that each EU donor shall not grant aid to more than three sectors in a given country. Furthermore, it explicitly defines different donor roles, with one of them being the

---

[64] For example, Dreher et al (2007) show empirically that the U.S. uses aid to buy voting compliance in the UN General Assembly.

[65] Even though donor coordination and donor specialisation generally are non-conflicting goals, Aldasoro et al. (2010) stress that these are two different concepts. Woods (2011) differentiates aid cooperation and aid coordination. The former concept entails that donors work together to plan, deliver and achieve common objectives, whereas aid coordination means that different activities are organised in a way that they do not impact negatively upon each other.

[66] See the following subsection for a detailed discussion.

'withdrawing donor' that will phase out its engagement in a sector (European Commission 2009).[67] The main purpose of such a specialisation is to improve aid effectiveness not only by mitigating the negative side effects of uncoordinated and fragmented aid by numerous donors but also by exploiting the respective comparative advantages of each donor in a country or sector (OECD 2011). The question is however, whether donors have an incentive to specialise on a certain set of partner countries, or whether they put less weight on aid effectiveness than on securing their global presence (Bigsten 2006). In fact, any specialisation (and also coordination among donors) effectively curtails their possibilities to pursue their commercial and political self-interest (Aldasoro et al. 2010).[68]

Even if a donor does not pursue its national interests when giving aid, negative spatial dependence might occur due to free-riding if poverty alleviation is seen as a public good. Economic development in a recipient country can be beneficial for all donors as it might improve international security by reducing the risk of terrorism or civil conflicts, but also as it creates potential export markets.[69] Since no donor can be excluded from these benefits, a single donor has an incentive to free-ride, i.e. to reduce its efforts in a recipient country if other donors provide aid.[70]

Consequences of spatial dependence

On the one hand, donor specialisation is not necessarily appropriate or a superior strategy under all circumstances. Especially after emergencies, such as the Tsunami hitting several Asian countries in 2004 or the recent earthquake in Haiti, great amounts of humanitarian aid are required that cannot be provided by a single country. Also, collective debt relief gives leeway to highly indebted countries and reduces the problem of free riding if debt relief by some creditors makes repayment of debt to other non-participating countries more likely. On the other hand, a lack of donor specialisation often goes hand in hand with a lack

---

[67] The EU Code of Conduct focuses on sectoral activities, but the argumentation easily extends to a country level.

[68] Using sector level data Aldasoro et al. (2010) show that most donors have not concentrated their aid and that the lack of coordination among donors is still prevalent. They conclude that "the gap between the words and the deeds of major donors appears to be as wide as ever" (Aldasoro et al. 2010: 936). Dollar and Levin (2006) find that aid became more selective in terms of democracy and economic governance over time, in a sense that donors specialise unilaterally to collaborate with a subgroup of potential partner countries. Yet, if this set is the same for all donors, this would lead to positive rather than negative spatial dependence.

[69] Most importantly, the public-good character leads to underprovision of foreign aid if donors do not take the positive externality into account (Torsvik 2005).

[70] Of course, this argumentation does not apply if a donor uses aid to secure its national interest since in this case a given donor has an incentive to increase its aid efforts to avoid losing influence.

of donor coordination in the recipient country. This is particularly true for herding behaviour where a donor simply copies the actions of other donors. Also, if ODA is allocated as a result of strategic interaction among donors, a tight coordination of the resulting aid projects is unlikely. Uncoordinated donor behaviour can have adverse effects on aid effectiveness. These can broadly be subsumed under two main arguments: the origin of aid darlings and increased aid volatility. While these are closely interrelated as the latter is partly a consequence of the former, both aspects have their own negative implications. The first set of problems is concerned with the fact that large amounts of aid are allocated to a certain country (or sector) which is accompanied by some negative effects, whereas other countries as much in need (or sectors as underdeveloped) are marginalised and do not receive sufficient financial resources. Cassen (1994) already notes that donors reveal some type of herding behaviour, which might be seen as evidence of 'aid fashion' in ODA allocation. On a country level, this leads to aid darlings and orphans, where the former receive much more aid than their economic and political situation would suggest, whereas the latter group receives substantially less. On a sector level, the focus on a few key areas, such as education or democratisation, detracts funding from others both across and within countries.[71] This jeopardises sustainable development progress in the neglected areas and potentially leads to an administrative overload in the favoured nations or sectors due to the mere amount of aid or increased donor fragmentation. One concern regarding the total amount of ODA inflow is decreasing returns on aid (Dudley and Montmarquette 1976, Lensink and White 2001). Limited absorptive capacities in the beneficiary country will also curtail the potential positive impact of financial transfers if local administrations are overstrained and if they lack the ability to manage aid flows efficiently (Berg 1997).

If other bilateral donors and multilateral aid organisations follow the allocation decisions of a given donor rather than specialising on a small group of partner countries, proliferation in the number of donors that are active in a country occurs as a consequence. This increased donor fragmentation might negatively affect the effectiveness of aid projects:[72] First, dealing with a large number of donors increases transaction costs as each donor has its own rules, procedures and reporting requirements.[73] Second, many foreign aid projects are characterised by large fixed costs and high returns to scale. These remain unexploited if

[71] Anecdotally, a financial director of an NGO interviewed by Reinhardt (2006) stated: "Kids are hot right now. Political dissidence isn't. That makes things very complicated for us."
[72] This is not to say that donor monopolies in a recipient country are a more favourable solution. This could also negatively affect aid effectiveness and would increase dependency of the recipient country.
[73] In Tanzania, for instance, the government has to write about 2,000 reports to numerous donors and to deal with more than 1,000 delegations per year (World Bank 2003).

each donor works on its own projects (Svensson 2008). Third, donor fragmentation impairs the recipient's financial ability and administrative capacity to govern. This is particularly true for project aid, where the donor works directly with or funds service providers such as schools or medical facilities. Furthermore, donors tend to support only investments and expect the recipient country to supply complementary inputs such as staff or maintenance (Svensson 2008). In addition donors poach for the most talented local experts and are able to attract them by paying salary supplements, which then will no longer be available to the domestic public or private sector (Arndt 2000). Redundancies in the bureaucracies of multiple donors aggravate this problem. Knack and Rahman (2008) find that larger donor fragmentation is associated with worse performance in terms of bureaucratic quality improvements. This is corroborated by Djankov et al. (2009) who find that recipients with highly fragmented donors grow slower than recipient collaborating with fewer donors. They argue that this can be partly explained by the increased corruption in the recipient country which is positively associated with the degree of donor fragmentation. Finally, if the number of donors in a recipient country is large and no single donor has a large share in the aid market, the accountability of each donor will be reduced and they might be less concerned that recurrent expenses caused by today's investment are secured or whether the projects are mutually consistent (Knack and Rahman 2008). Taken together, increased aid inflows and donor fragmentation caused by spatial contagion in aid allocation might lead to a situation, in which the "whole of aid in some sense [is] smaller than the sum of the parts" (Cassen 1994: 175).

The second major problem caused by herding in ODA granting is the increased volatility of aid flows if a change in aid allocation by one donor triggers changes in the same direction by other donors.[74] The beneficial potentials of aid can be neutralised by high volatility and unpredictability, which impedes planning and effective fiscal and monetary policy (Bulíř and Hamann 2003). Lensink and Morrisey (2000) and Arellano et al. (2009) find significant negative effects of aid volatility on aggregate growth and consumption. Several studies have identified macroeconomic determinants of growth that exhibit an increased volatility as a consequence of greater aid fluctuations, for instance inflation (Fielding and Mavrotas 2005), fiscal policy (Fatas and Mihov 2008), or real exchange rates (Edwards and Wijnbergen 1989). Agénor and Aizenman (2007) show that aid recipients may fall into a

[74] Bulíř and Hamann (2008) find that aid volatility has increased over time. This finding is supported by Kharas (2008) who argues that the costs of volatility are large and that herding behaviour aggravates collective volatility. Negative spatial dependence, on the other hand, has the potential to reduce volatility.

poverty trap if unstable sources of funding inhibit governments to invest in projects that require a steady flow of financial resources.

3.3. Literature review

The determinants of aid allocation have been a popular topic in empirical research on ODA. McGillivray and White (1993) provide an early literature review, a more recent overview can be found in Neumayer (2003). Contrary to the plethora of studies focussing on the different political and economic determinants of aid allocation, work on spatial dependency in ODA allocation is by far more limited. Even though there are a few studies which include aid by other donors in the regression models, the authors do not talk about spatial dependence – and it might be well assumed that they are not even aware that they are in fact analysing spatial dependence among donors. As a consequence, none of them has explicitly modelled this spatial contagion by using theoretically justified and more elaborated weighting matrices. More precisely, all existing studies solely include other donors' aggregate commitments as an explanatory variable, i.e. they implicitly use a non-row-standardised weighting matrix (or a row-standardised weighting matrix if the average ODA from other donors were used) that contains only ones for all donor-pairs, i.e. each donor gives the same attention to any other donor.[75] Furthermore, the definition of the dependent variable often differs from the definition of the aid transfers by other donors. In this sense, the models estimated are not purely spatial lag models which further corroborate the assumption that the authors do not consciously analyse spatial dependence.

Donors have a limited aid budget which needs to be distributed to recipients. This decision procedure can be modelled either as a one-step or a two-step process. A two-step process distinguishes the eligibility stage and the allocation stage, where the list of recipients is set up in the former stage whereas the actual amount of aid is determined in the latter stage. Since donors usually do not give assistance to every potential recipient country but only to a selected set of partner countries, the first step may contain a large number of countries that do not receive any aid, which leads to problems in the empirical analysis.[76] In contrast, both stages are combined and estimated simultaneously in a one-step procedure. Existing studies cannot only be classified according to how the decision process is modelled, but

---

[75] The ODA flows from other donors are usually net of the ODA of a particular donor. This is equivalent to setting all diagonal elements of the weighting matrix to zero.

[76] See the methodology part in Section 4 for a discussion and the statistical model generally used.

also whether they focus on a single donor country or whether the analysis encompasses a larger number of donors.

Looking at donor-specific studies first, McGillivray and Oczkowski (1992) estimate determinants of British ODA commitments. Their analysis, which uses a two-part model for estimation, is based on eight independent cross-sectional estimations for the years 1980 to 1987 and covers 93 developing countries. They find a positive effect of other DAC members' gross disbursements to the same recipient in year *t-1* in the allocation stage, but not in the eligibility stage. The elasticities in the second stage range from 0.25 to very high 1.36 for individual years, the pooled cross-sectional analysis over all years yields a coefficient estimate of 0.88. This indicates that a one percent increase in the gross aid disbursements of other donors is associated with a 0.88 percent increase in British aid commitments. Tarp et al. (1998) analyse determinants of Danish aid allocation decisions. Their dataset spans from 1960 to 1995 and covers 132 recipient countries. To control for aid given by other donors, the aggregate ODA disbursements from other bilateral and multilateral donors to a specific recipient country is included as a control variable and normalised by recipient population. Estimating a Heckman 2SLS model with a random-effects model in the second step, they find a positive and highly significant degree of spatial dependence both in the selection as well as in the allocation stage. In the latter, the effect size is 0.55, indicating that if the aggregate ODA disbursement by other donors to a recipient rises by one dollar per capita, Denmark increases its aid commitment to the same recipient by 55 cent. Differences between countries with and without former colonial ties in the factors that determine Spanish aid allocation are the focus of Vázquez's (2008) analysis, which covers the years from 1993 to 2005. The sample includes 104 partner countries in the eligibility stage and 25 nations without colonial ties and 20 with post-colonial links in the allocation stage. The variable included to capture spatial dependency is share of each recipient country at the global ODA (i.e. total aid disbursed by multilateral and bilateral donors), excluding Spanish aid. He finds evidence that countries that receive more aid from other sources are more likely to be on Spain's recipient list, and that in the allocation of the aid budget there is a focus on countries with a higher aid-dependency, but no spatial dependence effect in either group. In the eligibility stage, an increment of one percent in the share of global aid of a given country increases the odds for this nation of being on Spain's recipient list by nearly 41 percent. The author interprets the positive effect in the eligibility stage as a lack of donor coordination. Lastly, in a study on Italian foreign aid, Maurini and Settimo (2009) exploit the OECD DAC database and analyse absolute net disbursements from Italy to 156 recipients from 1983 to 2006. To test for the influence of other donors' aid, the aggregate disbursements of multilateral DAC members and other bilateral DAC member countries are included separately. A Tobit estimation reveals a positive and highly significant effect of other aid to the same recipient for all periods. Depending on model specification and time-period, the elasticities for other bilateral donors' aid disbursements are between 0.58 and 1.39. The coefficient of aid disbursements by multilateral donors is only significant in the period from 1991 to 1998 and from 1999 to 2006, with elasticity estimates of 0.40 and 0.44, respectively.

In contrast, to the best of my knowledge, only three studies have used a dyadic dataset encompassing various donors and recipients. Berthélemy and Tichit (2004) use a dataset covering 22 DAC donors and 137 recipient countries from 1980 to 1999. The total aid commitments per capita by other bilateral donors are included as a control variable. The estimation approach is a one-step random-effects Tobit model. Over the whole period and sample, there is no conclusive evidence that aid by other donors matters. The coefficient is positive and highly significant, unless controls for primary schooling enrolment and infant mortality are included. However, the effect size is negligible: if other donors on average increase their aid commitment to a given beneficiary by 100 USD, a given donor reacts with an increase in the range of 62 cent to one USD. Dividing the sample into three decadal subsets, the authors find a negative and statistically significant effect during the 1980s (coefficient -0.0129) and no effect during the 1990s. The difference between both periods, however, is positive and statistically significant at the one percent level, providing tentative evidence that dependence on other donors has increased. An analysis of differences between donors reveals a positive and highly significant effect of other bilateral donors' aid for the UK, the United States, Sweden, Germany and the Netherlands and a negative and significant impact on Belgian, Irish and Italian ODA.

Berthélemy (2006) examines donor-specific determinants of aid allocation using the same dataset. While the main focus of his analysis is to find differences in the importance of donor interest vis-à-vis recipient needs, aid per capita committed by other donors is also included as a control variable. For all donors, it is found that more assistance from other bilateral and multilateral donors increases the probability that a given recipient receives a positive amount of aid in the selection stage.[77] In the allocation equation, the effect of other

[77] The significant effect for other bilateral aid disappears if a random-effect Probit model rather than a standard Probit model is estimated.

donors' ODA is positive and significant in OLS and a Heckman 2SLS estimation indicating a complementary relationship between aid commitments from different donors. On average, a one percent increase in the per capita commitment of other bilateral donors is associated with a 0.28 percent increase in the commitment of a given donor. For multilateral assistance, this effect is 0.23 percent. However, once fixed effects are introduced, the effect of bilateral commitments becomes negative and significant and a one percent increase is accompanied by a 0.14 percent reduction in the commitment for a given dyad. Turning to donor-specific determinants, Berthélemy classifies donors into 'altruistic', 'moderately egoistic' and 'egoistic' donors according to effect size of trade relations between a donor and a recipient. In the cluster of moderately egoistic donors, he finds evidence that Germany, Canada, Belgium and Finland regard their assistance as complements to other bilateral donors' ODA, whereas Japan, United Kingdom and United States react with an aid reduction if other donors increase their aid commitments to a specific donor. Finally, all egoistic donors, i.e. Australia, France and Italy, regard their ODA commitments as substitutes for other donors' assistance. These results are partly in contrast to the ones of the previous study by Berthélemy and Tichit (2004), which indicates that results are sensitive to the estimation method chosen. The main difference between Berthélemy and Tichit (2004) and Berthélemy (2006) is that estimation in the latter study is based on fixed effects while the former analysis uses a random-effects Tobit model. Since the allocation of bilateral aid might be influenced by unobservable factors which can well be correlated with the explanatory variables, using fixed effects at least mitigates the resulting omitted variable bias. As will be discussed below, the introduction of fixed effects is even more crucial when analysing spatial dependence.

In a recent study, Claessens et al. (2009) exploit the OECD DAC dataset to analyse how aid allocation criteria changed over time. Unlike previous studies, they use net aid transfers rather than aid commitment or disbursements as the dependent variable. Using a dyadic panel dataset covering 22 donors and 147 recipients from 1970 to 2004, they estimate determinants of ODA disbursements for three sub-periods (1970-1989, 1990-1998, and 1999-2004). They include net aid transfers by other donors as a control variable; however it is not reported how this effect changes over time. Compared with other variables, they find very high effects for the aid transfers by other donors, but they are estimated very imprecisely, rendering them statistically insignificant in the main specification. Generally, some evidence for positive spatial dependence is found, but the effect is sensitive to sample selection. Table 17 summarises the main results for studies with a comparable econometric approach.

Table 17: Results of previous studies

Study	Donor(s)	Recipients	Study Period	Dependent variable	Methodology	Weighting Matrix	Variable measuring other donors' behaviour	Main Findings
McGillivray and Oczkowski (1992)	United Kingdom	93	1980-1987	ODA commitments	Two part model	Unitary (non row-standardised)	ODA disbursements of other DAC donors	no evidence for spatial dependence in eligibility stage; positive effect in allocation stage
Tarp et al. (1998)	Denmark	132	1960-1995	ODA commitments	Heckman 2SLS	Unitary (non row-standardised)	ODA disbursements of other donors	positive spatial dependence in both stages
Vázquez (2008)	Spain	104 (first stage)/ 45 (second stage)	1993-2005	ODA gross disbursements	Two part model		Share of recipient at global aid disbursements of other donors	positive spatial dependence in eligibility stage, but not in allocation stage
Maurini and Settimo (2009)	Italy	156	1983-2006	ODA net disbursements	Tobit	Unitary (non row-standardised)	ODA from other DAC donors	comprehensive evidence for positive spatial dependence
Berthélemy and Tichit (2004)	22	137	1980-1999	ODA commitments per capita	Tobit	Unitary (non row-standardised)	Aggregate ODA commitments of other donors per capita	some evidence for positive spatial dependence in full model, negative spatial dependence during 1980s, positive for Germany, Netherlands, Sweden, UK and USA, negative effect for Belgium, Ireland and Italy

Table 17: Results of previous studies (continued)

Study	Donor(s)	Recipients	Study Period	Dependent variable	Methodology	Weighting Matrix	Variable measuring other donors' behaviour	Main Findings
Berthélemy (2006)	22	137	1980-1999	ODA commitments	Two part model	Unitary (non row-standardised)	Aggregate ODA commitments of other bilateral donors per capita & multilateral ODA per capita	negative effect of ODA from other bilateral donors in fixed effects model; positive effect for multilateral ODA; positive effect for Germany, Canada, Belgium and Finland, negative effect for Australia, France, Italy, Japan, UK and USA
Claessens et al. (2009)	22	147	1970-2004	Net aid transfer	OLS, FE	Unitary (non row-standardised)	Aggregate net aid transfers by other donors per capita	some evidence for positive spatial dependence

A very different approach is taken by Frot and Santiso (2011) and this is also the only study that addresses spatial dependence explicitly. Departing from the observation that traders in bonds and equities show considerable herding behaviour, they apply methods developed in the finance literature to aid allocation decisions of bilateral donors. Basically, their herding measure is based on the share of donors that increase their aid to a particular recipient in a year and it is assumed that in the absence of herding, increases and decreases should be randomly distributed. Uniquely, the decision process is modelled in a way that donors first decide which recipients get *more* and which get *less* aid than in the previous period and then allocate aid in a second step. This is in contrast with all other two-step empirical studies in which the first step represents the decision *whether or not* a country receives aid at all and the actual amount is determined in the second step. In the empirical analysis by Frot and Santiso (2011), only the first step is examined. The data is taken from the OECD DAC dataset, which covers 48 years from 1960 to 2007. The dataset includes up to 5,171 recipient-years and encompasses activities of 60 different donors. Based on three year averages, they find a herding size of about 10 percent, indicating that if half of the allocation changes are increases and half are decreases, an average recipient experiences that 60 percent of its donors increase (or decrease) their allocations. Stated differently, this means that 50 percent of the beneficiaries see 60 percent of their donors decrease their aid compared to the previous period, while 50 percent see 60 percent of their donors increase their aid allocation. Differentiating between bilateral and multilateral organisations, the study finds no evidence for herding among multilateral donors; however, it is left open whether they react to allocation decisions taken by bilateral givers. An analysis of the determinants of herding reveals that a new policy adopted in the recipient country positively affects herding – after such a transition, a country receives funds from 20 percent more donors than the average beneficiary. A positive effect is also found for the occurrence of natural disasters, whereas a transition towards a more authoritarian regime is punished through a higher-than-average proportion of donors reducing their aid during the transition period. While the analysis explicitly models herding behaviour, the limitation on decreases and increases neglects differences in the size of the change, i.e. a 50 percent increase is treated as the same as a 0.5 percent increase. Even though they show that dropping small changes does not significantly affect the size of the herding measure, the absolute size of the change should play an important role when measuring the presence of spatial dependence.

Summing up, there is some evidence that the aid allocation of a particular donor is affected by allocation decisions of other donors, as most of the existing studies find a positive relationship between aggregate or average ODA allocated by other donors to a recipient and a given donor's aid to the same beneficiary. However, each of these studies, including the work by Frot and Santiso (2011), is based on the extremely simplifying assumption that every donor reacts to aid allocation of any other donor in the same way. This implies that, for instance, multilateral donors are influenced by bilateral donors in the same way as by other multilateral institutions. Furthermore, and even more severe, it means that important donors like the United States are affected in the same way by aid allocation decisions of small donors such as Luxembourg as small donors are influenced by important players in the global aid landscape. This assumption is not adequate if a particular donor observes the aid allocation of some other donors more closely than the actions of others. As will be shown in the empirical part below, this assumption does not hold true – smaller donors are more influenced by larger donors than vice versa. Furthermore, all existing studies imply that the strength of this spatial dependence does not depend on the strength of the relations between a donor and recipient. The present analysis is the first to model the donor-recipient relationship explicitly. This also allows testing for strategic interaction among donors, i.e. if a donor is likely to increase its aid allocation to an important partner country, if other donors with also close ties with the same recipient increase their aid efforts in this country. This is particularly important as beneficiaries differ with respect to their geo-strategic position, but also with respect to their market potential or their endowment with natural resources. If donors use aid strategically to gain influence in the recipient country, these potential benefits in the receiving country should impact the strength of spatial dependence and therefore should be modelled directly.

3.4. Methodology and Data Description

The following section first describes the dataset used and defines all variables employed in some detail. Then, the rationale behind the specific operationalisation of the spatial lags and their calculation is discussed, while the last subsection is concerned with the model specification and the estimation technique.

Data description

The OECD defines official development assistance (ODA) as grants or loans to developing countries that are undertaken by the official sector. These funds are given mainly to promote economic development and welfare, and contain a grant element of at least 25 percent in the case of loans (OECD 2008).[78] This definition includes administrative expenses in the donor countries, technical assistance and financial flows. For the present analysis, the aid definition is further restricted to country programmable aid (CPA). CPA is defined as total gross ODA net of aid that is (1) unpredictable by nature (2) entails no cross-border flows, and (3) is not part of co-operation agreements between governments (OECD 2009). As a consequence, all projects classified as humanitarian aid, debt relief, food aid and administrative costs in the donor country are dropped from the sample. This leaves core aid that finances development in a medium to long term perspective and is closer than ODA to representing aid flows that are relevant to the decision-making at donor level (OECD 2009). Furthermore, high debt levels, famines as a consequence of crop failure or the need for humanitarian aid after natural disasters in a certain country may trigger increased aid allocation by several donors to this recipient. This behaviour does not need to reflect spatial dependence, but it could be purely driven by a phenomenon called spatial clustering.[79] Therefore, excluding aid projects that are particularly prone to spatial clustering facilitates interpretation as spatial dependence.

Data on aid is taken from the Project Level Aid (PLAID) database provided by the College of William and Marry, Brigham University, and Development Gateway (Nielson et al. 2010). The PLAID contains information on aid projects not only by members of the Development Assistance Committee (DAC), but adds activities by non-DAC multilateral organisations. Furthermore, it provides more detailed project descriptions, which allows a

[78] To restrict the sample to ODA, several steps are undertaken: First, all observations with missing information on both the flow type and the grant element are deleted. Then all projects classified as loans at market rates, equity investment, unknown loan type or other official flows are excluded. Finally, all observations with a grant element of less than 25 percent are dropped. Unfortunately, several multilateral organisations, among them the World Bank, do not report the grant element of their commitments and are therefore excluded.

[79] Spatial clustering refers to the fact that close units are more similar than distant units due to observed characteristics, while unobserved determinants of spatial patterns are denoted unobserved spatial heterogeneity (Plümper and Neumayer 2010). In the context of aid allocation, increased aid by several donors due to natural disasters is an example of spatial clustering, whereas less aid due to a general bad reputation of a recipient country is unobserved spatial heterogeneity. The problem of separating the influence of spatial clustering and unobserved spatial heterogeneity and identifying spatial dependence is commonly referred to as Galton's Problem (Galton 1889).

more precise allocation of projects to sectors.[80] The dataset encompasses information both on commitments and disbursements from 1945 to 2009; however, with the exception of a few donors such as the World Bank, the vast majority of donors start reporting in 1974 or later. Since the purpose of this study is to examine how different donors depend on each other, information on a large number of donors is crucial in order to avoid biased results. Therefore, the study period is restricted to 1974 to 2008, which contains over 99 percent of all observations. In order to keep the sample at a manageable size, several alterations are made: On the donor side, multilateral organisations are replaced by their umbrella organisation and minor organisations with few projects or a very constrained field of activity are deleted. In addition, smaller donors with less than 1,000 projects over the whole study period are neglected.[81] Regarding the recipient-sample, projects with missing information in this variable are dropped, as are observations for which no specific country is listed as a recipient. This concerns particularly multilateral organisations and regionally unspecified flows. Additionally, OECD members as recipients are deleted as they are not the main target for aid. All recipients with fewer than 100 projects and/or a population of less than 100,000 people are disregarded. Finally, ten recipients are dropped due to missing data on the control variables. This leaves a sample of 23 country donors and three multilateral institutions on the one hand and 139 recipient countries on the other. Tables 27 and 28 in the Appendix list all countries and institutions included.

ODA commitments rather than the actual disbursements are taken as the dependent variable.[82] As argued by Dudley and Montmarquette (1976), these provide a more accurate measure of donor supply than the actual money transfers, which partly depend on the administrative capacity and willingness to accept the funds in the recipient countries (Berthélemy 2006).[83] After dropping a very few number of projects with commitments of zero or less, the data are collapsed to form a donor-recipient-year dataset. Assuming full reporting, all gaps in the dataset were set to zero, but replaced by missings for all donor-years in which a particular donor does not report any figures. Also, for recipients that became independent only after 1974 (mainly members of the former USSR), all

[80] The PLAID dataset also extends coverage to bilateral non-DAC members; however, these are not included it this analysis due to their small number of projects or due to missing information on the loan type or grant element.

[81] 16 donors are affected, which are mainly developing or small countries such as Monaco, Iceland, and Liechtenstein.

[82] The OECD defines commitments as “a firm obligation expressed in writing”, while disbursements are “the actual transfer of financial resources” (OECD 2002: 292).

[83] Besides this theoretical reasoning, with 50 percent missings information on disbursement is far less complete in the PLAID dataset.

observations before the independence year are set to missing.[84] After cleaning up the dataset, it contains information from 600,376 projects and covers commitments of 1.14 trillion USD (in constant 2000 USD).[85] There is no reason to assume that the necessary sample reductions are non-random and cause any sample-selection-bias.

Most existing studies use either total amount of aid or aid per capita as left-hand variable; advocates of the former argue that the allocation on a per-capita basis "is both a difficult and cumbersome task" for donors (McGillivray and Oczkowski 1992: 1314). On the other hand, econometricians using per capita aid stress that it automatically controls for country size and allows testing for a small-country-bias (Berthélemy and Tichit 2004). However, a donor usually has a fixed aid budget, which is divided among potential recipients. This decision is best approximated if the aid to a particular recipient is expressed as a share of the total aid allocated by a donor in a given year (Neumayer 2003). This definition has two other virtues (Vázquez 2008): First, it eliminates any bias caused by comparing figures over different years, for instance caused by measurement errors due to fluctuations in domestic exchange rates to the USD. Furthermore, it allows the use of nominal commitments and renders the choice of the correct deflation factor unnecessary. Second, as shares are normalised to one per donor and year, it is insensitive to trends in the size of the aid budget over time, e.g. the widespread reduction of aid budgets in the 1990s.

The set of explanatory variables can broadly be divided into three subsets: the spatial lags, which are employed to test for the existence of spatial dependence and are described in more detail in the following subsection, variables that measure donor interest and controls for recipient need. The set of variables that control for recipient need encompasses:

- The *GDP per capita* in constant 2000 USD taken from the World Development Indicators (World Bank 2010) is included as an approximation for different aspects of development of the recipient country. It is expected that less aid is allocated to richer countries.
- To control for the size of a recipient, its *Population* is taken from the World Bank (2010). All else being equal, larger countries should receive a larger share of the aid budget of a donor.

[84] In some cases, aid flows are reported to dependent countries. Here, all years prior to the first positive ODA flow are set to missing.
[85] This analysis is based on the 1.9.2 version of the dataset.

- The number of fatalities caused by natural disasters (*Disaster deaths*) is taken from by the Emergency Events Database (EM-DAT 2010) and is expressed as number of deaths per 1,000 people of population to assess the impact on a country. Even though humanitarian aid is excluded in the aid definition used, Frot and Santiso (2011) argue that natural disasters draw attention to the affected country and trigger aid flows from several donors not necessarily only in form of disaster relief but also in form of other long term investments. Accordingly, it is expected that a greater share of ODA is directed to countries that are hit by natural catastrophes.
- As a proxy for democratisation and good governance (*Democracy*), a measure provided by the Freedom House (2009) is included. As argued in the introductory section, a higher degree of democracy should raise the effectiveness of aid and send a positive signal to donors, thus a positive sign of the coefficient is anticipated.

With regard to donor interest, the following set of variables is taken:

- To account for economic interest, the variable *Bilateral trade* measures the trade between the donor and the recipient, where trade is defined as the sum of imports and export. Bilateral trade is divided by donor's GDP to reflect the importance of the trade relations for the donor. Trade (and particularly exports) exhibits the risk for a potential degree of simultaneity, in case tied aid increases the bilateral movement of goods. However, since commitments rather than disbursements are used in this context, and as the latter lag behind the former, the risk should be limited (Berthélemy 2006).
- As a measure for the strategic importance of the recipient, its share in the total US military grants is taken (*US military grant share*). Data is provided by USAID (2010). Unfortunately, this information is not sufficiently available for other donors. The underlying rationale is that countries that receive a large share of US military grants are strategically important to Western donors. It is expected that also more aid resources are directed to these countries.
- As an approximation for the bilateral political relations, a dummy that takes the value of one if either of the dyad member has dispatched an ambassador to the other member dyad (*Diplomatic representation).* As more assistance should be allocated to countries with which the donor has close political ties, a positive effect is anticipated.
- To account for similarity of state preferences between donor and recipient, the variable *UN voting similarity* is included. It measures the similarity of the voting

positions of donor and recipient in the United Nations General Assembly and is taken from Gartzke (2010). The variable is recoded to run from 0 (complete different voting positions) to 2 (always the same voting position). Again, a positive sign is anticipated.

- A dummy for the existence of a *Military alliance* is included. This variable takes the value of one if the donor and the recipient have signed a defence or an offence pact. The data is provided by Leeds et al. (2002). One argument for giving foreign aid is provided by strategic rationales; therefore, a positive coefficient is anticipated.
- Even though the dependent variable is the share of aid of a given recipient at the total aid budget of a donor and this measure is independent from the overall size of the aid budget, a *t-1* set of year dummies is included in order to capture any general time trend in aid allocation decisions which is induced by changes in the set of potential recipient countries due to data availability or political transformation. Most importantly, the members of the former Soviet Union emerged on the list in 1990, but also country splits affect the sample.

To include bilateral as well as multilateral organisations in the sample, the voting shares of the member states are used as weights to determine bilateral relations between a multilateral organisation and the recipients for the variables for donor interest and variables that link a donor with a recipient (e.g., distance and colonial ties).[86] This reflects the assumption that in multinational donor institutions the power of a single member state to assert its national interest is determined by its relative influence in votes.[87] With the exception of disaster fatalities, all time variant variables are lagged by one year to mirror the situation allocators of aid faced at the time of decision-making and to reduce the potential risk of endogeneity.

There are several reasons to assume that aid commitments are time dependent, i.e. that the amount of the aid committed in a year depends on the assistance given in the year before. One of them is bureaucratic inertia: aid allocation decisions are the result of bureaucratic procedures, which allow only incremental change as long as there is no pivotal intervention (Allison 1971). Giving aid to a constant set of recipients also minimises administrative

[86] Due to restricted data availability, all voting share information is taken as time invariant.
[87] This is also true for dummy variables (e.g., diplomatic representation, military alliance). Thus, they are not dummies any more in the final dataset. However, interpretation is not affected.

costs and efforts, as adding new partner countries involves additional expenses to implement new bilateral mechanisms. Furthermore, administrative efficiency might be increased if learning economies can be realised based on previous experiences in a recipient country (Vázquez 2008). As a donor is likely to be better informed about the economic and political situation as well as the need in a particular country in case this country is already a recipient of bilateral aid, new aid commitments to this country exhibit less risk than aid to a new beneficiary. By the same token, experiences with the effectiveness of previous aid render it easier for a donor to assess ex-ante the outcome of a potential aid project. Particular dyad specific characteristics, for instance between Spain and many Latin American countries, may also lead to a cooperation which exhibits an increased stability over time. Lastly, sustainable development is a long-term process so that a donor might provide stable assistance over time to a partner as long as the terms of the cooperation are fulfilled. To account for this time dependency and to reduce a potential bias due to temporal dynamics, a one-year lag of aid commitments is included as a control variable.

Calculation of spatial lags

To analyse the potential influence of aid allocated to the same recipient by other donors, a spatial lag model is used. In such a model, for each observation the dependent variable of all other observations is included as a right-hand side variable. This variable is weighted using a weighting matrix. These matrices are an integral part of spatial modelling and formally delineate the spatial dependence between a given observation and all other observations (Anselin 1988). Aid flows from a donor to a recipient, either bilateral or multilateral, are an example of a directed dyad, in which there is a clear source and a target and the action originates at the former and is directed towards the latter. As outlined by Neumayer and Plümper (2010), there are various options to model spatial dependence in such a setting. In the present context, spatial dependence is assumed to take the form of 'specific source contagion' (Neumayer and Plümper 2010: 154), in which the aid flow (or the probability of a positive amount of aid in the first stage) between a donor *i* and a recipient *j* depends on the aid flows (and the existence of a positive amount of aid, respectively) of other donors *k* with the very same recipient *j*. For example, the amount of aid the United Kingdom gives to Ghana depends on the assistance that other donors (e.g., France and Germany) allocate to Ghana. There are other forms of spatial contagion, such as aggregate source or aggregate target contagion, and specific target contagion (Neumayer and Plümper 2010). The aggregate forms of contagion assume that aid flows between all other dyads, and not only those including recipient *j*, influence the aid from *i* to *j*. Against the background of donor interests in the recipient country it is not reasonable to presume that aid flows by other donors to other recipients influence aid allocation to recipient *j*. Specific target contagion assumes that the amount of aid a recipient *j* receives from a donor *i* spatially depends on the amount of aid this donor *i* gives to other recipients *m*. Given the fixed aid budget, aid has to be allocated between all potential recipients. This automatically introduces dependency between the amount of aid given to recipient *j* and other recipients *m*. In the present analysis, this is captured by expressing aid to recipient *j* as a share of total aid by donor *i*. Other than the fixed aid budget that needs to be allocated, there are further arguments why such a specific target contagion might exist also in absolute aid flows. For instance, one recipient might successfully combat corruption and thereby improve its relative governance-performance among all potential recipients. If a donor rewards this improvement with more aid to this country, there would be spatial dependence between recipients. Since the focus of this work is spatial dependence among donors, however, specific target contagion is not modelled explicitly.

The spatial lags are the product of the dependent variable of all other $dyads_{kj}$ and the weighting matrix, which expressed the connectivity between $dyad_{ij}$ and $dyad_{kj}$. Abstracting from all other explanatory variables, the model reads as follows (Neumayer and Plümper 2010):

$$Y_{ij} = \rho \sum_{k \neq i} w_{ik} Y_{kj} + \varepsilon_{ij} \qquad (11)$$

where Y_{ij} is the foreign aid of donor *i* to recipient *j*, Y_{kj} the aid of other donors except donor *i* to the same recipient *j* and w_{ik} is the weighting matrix which measures the connectivity between donor *i* and donor *k*.[88] The parameter ρ is the spatial lag coefficient to be estimated. The modelling of the strength of spatial dependence between two donors deserves closer attention. As argued in section 2, it is assumed that donors take the aid decisions of other donors into account when allocating their aid. The basic questions are (1) whether all other donors *k* have the same influence on donor *i* or whether there are some donors that exert more influence than others, and if so, (2) which donors are more influential. To test this, six different weighting matrices are used to construct the spatial lags, which can be grouped into two categories: the first group encompasses global

[88] For illustrative purposes, time indices not included.

weighting matrices that are not recipient specific,[89] whereas weights in the second group depend both on the link between donor *i* and recipient *j* as well as donor *k* and recipient *j*. The first three weighting matrices, which do not depend on the recipient, are used to test for the existence of a general herding behaviour among donors, while the latter three weighting matrices model strategic interaction among donors.

Specifically, the three global weighting matrices in the first group are:

- *All donors*: This is a very basic unit matrix, which consists of only ones for all non-diagonal elements.[90] This implies that all donors exert the same degree of influence. For instance the impact of the dyad Germany-Ghana on UK-Ghana is the same as the impact of the dyad France-Ghana. Such a weighing matrix simply measures the aggregate share of a given recipient country at the aid from other donors (if matrix is not row-standardised) and the average share of aid, respectively, with a row-standardised matrix.
- *Global aid share*: In this weighting matrix, the cells contain the total amount of CPA of a donor in a given year. As in the first weighting matrix, all rows of the matrix contain the same values (with the exception of diagonal elements which are set to zero), because the total amount of CPA of one donor does not depend on who the corresponding other donor in the matrix is. After row-standardisation, each cell contains the share of a given donor (except donor *i*) at the total amount of CPA given by all donors (except donor *i*) in a year and measures the relative importance of a donor. This matrix reflects the assumption that a given donor does not follow each other donor in the same way, but that it gives more weight to the aid allocation decisions of important donors. For example, UK's aid allocation is assumed to be more affected by the aid giving of the United States than by Luxembourg's aid allocation.
- *Same subgroup*: Here, donors are classified into four groups and donors are assumed to be impacted only by other donors in the same subgroup. The classification of donors is according to Neumayer (2003): The group of Big Western donors encompasses six traditional donors with large (absolute) aid budgets that account for 60.3 percent of all CPA in 2008. The like-minded countries are a group of bilateral donors who have the reputation that the promotion of

[89] For these, the weighting matrix w_{ik} is equal to w_{ki}.

[90] The diagonal elements of all weighting matrices are set to zero since a donor cannot spatially depend on itself.

democracy and good governance plays a particularly important role in their aid allocation decisions and that their aid is more directed towards poverty alleviation (Stokke 1989). Finally, other bilateral donors are all remaining country donors in the sample who are not part of the former two groups and multilateral donors are multilateral institutions (see Table 27 for a classification of donors).

In the second set of weighting matrices, the influence of donor *k* on donor *i* regarding aid to recipient *j* depends on the link between the donor *k* and recipient *j* on the one hand and the link between donor *i* and recipient *j* on the other hand. These links can be modelled as substitutes or as complements. The former assumes that the strength of connectivity between $dyad_{ij}$ and $dyad_{kj}$ is positive even if only one of the links (i.e. between *i* and *j or* between *k* and *j*) is non-zero, whereas for a complementary relationship both links must be positive. Mathematically, substitutes are expressed by the sum of the two links, and complements by their product. In the present analysis, a complementary relationship is assumed, that is, $dyad_{ij}$ is influenced by $dyad_{kj}$ if, and only if, both the $link_{ij}$ *and* the $link_{kj}$ are non-zero:

$$w_{ik} = w_{(ij)(kj)} = link_{ij} \times link_{kj} \quad (12)$$

The weighting matrix, which determines the influence of aid flows from donor *k* to recipient *j* on the aid flow from donor *i* to recipient *j* is the product of the link between donor *i* and recipient *j* and donor *k* and recipient *j*. This reflects the assumption that aid to a given recipient is only influenced by aid from other donors to the same recipient if this recipient is of importance for both donors: e.g., for economic considerations that determine aid allocation, aid by donor *i* to recipient *j* is the stronger influenced by aid from donor *k* to the same recipient *j* the more economically important recipient *j* is for donor *i* and the more economically important recipient *j* is for donor *k*. Economic importance is here measured by a recipient's share of the total trade of a donor.

To represent the different aspects of donor interest discussed above, three different measurements are taken as a link:

- *Bilateral trade*: To account for economic interest, the product of the two absolute trade values (sum of bilateral imports and exports) is taken. $link_{ij}$ is the total trade value of recipient *j* with donor *i*, whereas $link_{kj}$ is the value of trade between recipient *j* and donor *k*. The value in the respective cell of the weighting matrix is the higher, the more both donor *i* and donor *k* trade with recipient *j*. For instance, China is a relatively important trade partner of Japan and South Korea. Therefore, it is assumed that the aid allocation decision of Japan with regard to China is relatively strongly influenced by the aid from South Korea to China.

- *Military alliance*: This weighting matrix is used to test the hypothesis that a donor is more heavily influenced by other donors if a particular recipient is strategically important for both donors. As described above, the variable is coded as one if both countries of a dyad formed a military alliance and zero otherwise. The weighting matrix contains the value of one, if and only if, both donor i and donor k have a military alliance with recipient j.
- *UN voting similarity*: This measure is a proxy for close political relations. As with bilateral trade, the weighting matrix is the product of the two links between donor i and recipient k and between donor j and recipient k. The link is similarity of the voting behaviour in the UN General Assembly (see the variable description for details).

For ease of interpretation, all weighting matrices are row-standardised, i.e. each cell of the matrix is divided by its row sum. This results in a new row-standardised weighting matrix in which the weights in each row add up to one. The spatial lag is now the weighted average of the spatially lagged dependent variable in other dyads rather than its weighted sum. Even though row-standardisation is standard in the spatial econometric literature, Plümper and Neumayer (2010) argue that it must be well justified as it changes the relative influence of other units and may thereby alter estimation results. Here, the spatially lagged variable is the share of aid of another donor to the same recipient rather than the total amount of aid given by this donor to the same recipient. Since the sum of shares has no particularly meaningful interpretation in any case, row-standardisation is adequate from this perspective. However, for the three global weighting matrices, row-standardisation does not affect the results: since the rows of the weighting matrix *Global aid share* by definition already sum up to one, the matrix is implicitly row-standardised. For the weighting matrices *All donors* and *Same subgroup*, which give equal weight to either all or a subset of other donors, the coefficients will change depending on whether the weighting matrix is row-standardised or not, but the significance levels are not affected.[91] For the three recipient-specific weighting matrices, row-standardisation does affect the results. The

[91] These weighting matrices contain a fixed number of ones in each row: For the weighting matrix *All donors*, the number of ones is 25 for the 25 other donors in the sample. For the spatial lag with the *All donors* weighting matrix, the effect of the spatial lag with a non row-standardised weighting matrix is simply the effect of the spatial lag using a row-standardised matrix times 25.

argument for doing so is that one is not interested in the absolute values of the elements in the weighting. Rather, the elements in the weighting matrix determine the relative impact of other donors. The row-standardised *Bilateral trade* weighting matrix allocates weights according to the relative importance of a recipient country for the two donors *i* and *k* and should not be driven by the fact that larger countries generally trade more than smaller countries. The products of the bilateral trade volumes between donor *i* and recipient *j* and donor *k* and recipient *j* are summed up for each donor *i*. If the weighting matrix is not row-standardised this would assign a stronger influence to countries with higher absolute trade values as for these countries the links are higher.[92] If the weighting-matrix *Military alliance* is not row-standardised, the effect of an increased number of military alliances would also be captured in the spatial lag. However, the focus is not on the number of such alliances concluded. Similarly, one is neither interested in the aggregate similarity in the UN general assembly. Besides this theoretical reasoning, row-standardisation provides two other virtues as noted by Ward and Gleditsch (2008). First, the spatial lag has the same unit as the dependent variable (here percent of aid allocated to a given recipient), which facilitates interpretation. Second, row-standardisation allows for directly comparing the effect of a temporally lagged and a spatially lagged dependent variable. Since all estimations include a one-year lag of the dependent variable, this provides a further reason for row-standardising the weighting matrices.

Table 18 presents the summary statistics. While in the whole sample, 51.6 percent of all dyads have a positive amount of aid, in the sample used for the fixed effects first stage estimation, the share is 54.5 percent.[93] In the total sample, the average share of a recipient at the total amount of aid of a donor in a year is one percent. If the sample is confined to the 30,396 dyads with a positive amount of foreign aid, this mean increases to 1.92 percent.

---

[92] Furthermore, since the elements in the weighting matrix are the product of the trade volumes of donor i and donor k, the values of the spatially lagged dependent variable would be higher for a donor with high trade values. There is no reason to assume that countries which trade a lot are generally subject to a higher degree of spatial dependence. The same argument applies to the other two recipient-specific weighting matrices.

[93] As discussed below, the reason for this is that under a fixed-effects approach, all dyads with either a positive amount of aid for all years or no aid in any year are dropped due to a lack of overtime-variation. The fact that the share of dyads with a positive amount of aid is higher in the fixed effects sample than in the total sample indicates that there are more dyads with no positive amount of aid in any year than dyads with a positive amount of aid throughout.

Table 18: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Dependent variables					
CPA share	59,030	0.990	3.638	0	100
CPA dummy	45,691	0.545	0.498	0	1
Control variables					
Disaster deaths	59,030	0.020	0.234	0	8.044
Democracy (t-1)	59,030	6.855	3.334	2	14
US military grant share (t-1)	59,030	0.864	4.889	0	59.476
Bilateral trade (t-1)	59,030	0.079	0.209	0	5.293
GDP per capita (ln, t-1)	59,030	6.830	1.330	4.131	10.806
Population (ln, t-1)	59,030	15.914	1.727	10.600	21.000
Military alliance (t-1)	59,030	0.058	0.220	0	1
Diplomatic representation (t-1)	59,030	0.650	0.445	0	1
UN voting similarity (t-1)	59,030	1.345	0.286	0	2
Spatial lags					
W: All donors (t-1)	59,030	0.786	1.149	0	9.082
W: Global aid share (t-1)	59,030	0.968	1.640	0	15.627
W: Same subgroup of donors (t-1)	59,030	0.833	1.803	0	42.824
W: Bilateral trade link (t-1)	59,030	1.731	3.291	0	46.721
W: Military alliance link (t-1)	59,030	0.081	0.792	0	42.663
W: UN voting similarity link (t-1)	59,030	0.843	1.229	0	12.005

Notes: W: denotes the weighting matrix used for calculating the spatial lags; with exception of the variable *CPA dummy,* values of all other variables relate to the second stage sample without the restriction to positive values of the dependent variable.

Model specification and estimation methodology

As noted by Plümper and Neumayer (2010), model specification in the analysis of spatial dependence needs to take into account several specific issues. This is to avoid biased results and to draw causal inference rather than simply catching spurious effects. First, the one-period time lag of the dependent variable that is included on the right hand side controls for temporal dynamics. The effect of a common trend in the size of the aid budget, e.g., all donors give more or less aid, is removed by normalising the aid commitments per donor-year to one, i.e. by expressing aid in shares rather than in absolute values. A *t-1* set of year dummies controls for a change in the number of recipients for a given aid budget which would lead to higher or lower average shares for all recipient countries. Furthermore, the existence of spatial clustering and unobserved spatial heterogeneity, i.e. factors which influence aid allocation decisions of several donors in the same direction but which cannot be controlled for can lead to biased spatial effects. To mitigate the impact of the former, the model specification is as broad as possible to control for a wide range of observable factors that might influence donor decision.[94] To address the problem of unobserved spatial heterogeneity, all models are estimated with dyad fixed effects. This removes all variation between dyads and the estimation is solely based on the with-in variation of each dyad. While this automatically controls for any time-invariant dyad specific effect, such as cultural and geographic proximity or bilateral relations (for example the United States' large aid to Israel and Egypt), it also removes unobserved spatial heterogeneity and spatial clustering in aid levels (Plümper and Neumayer 2010). The inclusion of fixed effects also changes the tested hypothesis: instead of examining spatial dependence in the amount of aid, now spatial dependence in the changes in aid shares over time is analysed. In the present context, this is reasonable and more in line with Frot and Santiso (2011). Spatially lagged dependent variables exhibit a certain degree of endogeneity as an external shock that cannot be controlled for in a given dyad affects other dyads via the spatial lag and is reflected back to the original dyad on the same way. However, this bias should be less pronounced in aid shares than in aid levels. More importantly, the spatial lags are lagged by one year which further mitigates the problem.

The process of aid allocation can be modelled as a two-step decision: In the first step, a donor country decides to which of all potential recipients it will allocate any positive amount of aid (eligibility stage, gate-keeping state). In case of a positive answer, the actual amount of aid is determined in a second step (level stage).[95] Since many donors give some foreign aid to selected countries and nothing to others, the dependent variable is zero in many cases. Major donors thereby tend to disburse their aid more widespread, whereas smaller donors tend to concentrate their aid on a few beneficiaries (Isopi and Mavrotas 2006). Thus, the dependent variable is only partly continuous and has a positive probability mass at the value of zero, which violates the OLS assumption that the expected value of the dependent variable is linear in the explanatory variables. Neumayer (2003) provides an overview over the three most commonly used econometric approaches to deal with this

[94] However, there is a clear trade-off between the comprehensiveness of the model specification on the one hand, and the problem of data availability and multi-collinearity on the other.
[95] Vázquez (2008) proposes a three-step model, in which the first step is the decision of a government on the size of the ODA budget. Also this stage can be subject to spatial contagion if the amount dedicated to ODA in one country depends on the aid budget in other donor countries. However, as this decision is affected by a completely different set of determinants, it is ignored in this analysis. See e.g. Hopkins (2000) and Round and Odedokun (2004) for studies.

situation: the Tobit model, the Heckman sample selection model, and the two-part model. The Tobit model, on the one hand, is a one step estimator. Therefore, it assumes that the set of variables that determines whether a country receives any aid (eligibility stage) is the same as the set influencing the decision how much aid is allocated in the second stage. It is also constrained in the sense that it does not allow for the fact that a certain factor increases the probability of a country to be eligible, but decreases the actual amount of aid allocated (i.e. the sign of the coefficient is the same in both stages). While these assumptions are possibly plausible, they are quite restrictive and stipulate effects which should be tested rather than be taken as given. In the Heckman sample selection model, on the other hand, the eligibility stage is estimated as a Probit model, in which a binary dependent variable takes the value of one if a recipient receives any positive amount of aid in a given year and zero otherwise. To correct for the sample selection bias due to the endogeneity in the selection process, the inverse Mills-ratio obtained in the first step is then included as an explanatory variable in the second step (Berthélemy 2006). The most important caveat of this approach is that one needs to find a so called exclusion restriction for reliable results, that is a variable that only affects aid eligibility, but has no influence on the actual amount of assistance given (Neumayer 2003). Such a variable is usually hard to find since factors influencing the probability that a country is a recipient of aid (e.g., poverty, population or strategic importance) also affect how much aid is given. Against this background, the third alternative, namely the two-part model, is the preferred estimation technique in this analysis. It was first introduced by Cragg (1971) and has been widely applied in the context of aid allocation. It resembles the Heckman sample selection model in also estimating two separate equations for both stages, but is based on the assumption that the two stages are independent of each other, i.e. that there is no correlation in the error terms of both regressions.[96] When analysing aid allocation, this might be a reasonable supposition, since the first step, i.e. the decision which countries will be partner countries, is more political and regularly influenced by the government (as pointed out by Tarp et al. (1998) for the Danish aid policy), whereas the second step, the actual allocation is determined by the aid

[96] A test on the significance of the correlation parameter in the Heckman sample selection estimations can reveal whether this assumption is appropriate. However, as shown by Manning et al. (1979), the potential bias of the two-part model is likely to be minor in typical situations. In the context of ODA allocation, Alesina and Dollar (2000) and Berthélemy (2006) do not find much correlation between the residuals of the selection equation in the first step and of the allocation equation in the second step. They conclude that the linear estimation in the second step is as good as the Heckman estimation.

administration. Eventually, the model estimating the eligibility stage reads as follows and is estimated with a conditional fixed-effects Logit model:[97]

$$\Pr(Y_{ijt} = 1) = F(\alpha_0 + \beta_1 Y_{ij(t-1)} + \rho' X_{ijt} + \pi' W_{ijt} + \delta' T_t) \qquad (13)$$

where Y_{ijt} is a dummy taking the value of one if recipient *j* receives any positive amount from donor *i* in period *t*.[98] $Y_{ij(t-1)}$ is the one-year lag and indicates whether the same recipient received aid from the same donor in the prior period. The spatial lags described above are included in X_{ijt}. W_{ijt} contains the control variables. Finally, *T* is a *t-1* set of period dummies to capture aggregate effects such as the total amount of aid allocated. *F* stands for the cumulative standard normal distribution. The parameters *α*, *β*, *ρ*, *π*, and *δ* are to be estimated. The estimation equation for the second stage is as follows:

$$Z^*_{ijt} = \alpha_0 + \beta_1 Z_{ij(t-1)} + \rho' X_{ijt} + \pi' W_{ijt} + \varepsilon_{ijt} \qquad (14)$$

where Z_{ijt} is the share of recipient *j* at the total CPA of donor *i* in period *t*.[99] Since in the allocation stage the actual amount of aid given is estimated, only observations with $Z^*_{ijt} = Z_{ijt} > 0$ are included in the analysis, while dyads with no positive aid commitment in year *t* are dropped. $Z_{ij(t-1)}$ again is the one-year lag of the dependent variable and ε_{ijt} is the error term. In both stages, the main interest of the analysis lies in the coefficient of the spatial lags *ρ*. Estimating the dynamic model in equation (4) with a fixed effects model introduces a Nickell (1981) bias; however this bias diminishes as the number of periods *T* gets large and the dataset covers the period from 1974 to 2008.

---

[97] This deviates from other studies which use a Probit model to estimate the first stage (e.g., Dudley and Montmarquette (1976), Svensson (1999), Neumayer (2003)). While the results of a Logit and a Probit estimation are very similar in general, a conditional fixed-effects model is not available for the Probit estimator. However, as argued above, dyad fixed effects are necessary to remove spatial clustering in levels and render the interpretation as causal effects more likely. Katz (2001) shows that there is no bias in the conditional fixed-effects Logit model even if T is smaller than 20.

[98] Some others, such as McGillivray and Oczkowski (1992) and Vázquez (2008) define a country as an aid recipient only if the aid received exceeds a given threshold, e.g. one percent of Spanish ODA in Vázquez (2008). However, the choice of this cut-off point is arbitrary as there is no guiding theory.

[99] Formally, this can be expressed as $Z_{ijt} = (\frac{CPA_{ijt}}{\sum_j CPA_{it}}) \times 100$, $Z_{ijt} \in [0;100]$.

3.5. Main Results

The presentation of the results is organised as follows: First, the estimations for the eligibility stage are shown and discussed for all donors and the whole study period, followed by the results for the allocation stage for the same sample. Since the donor landscape is heterogeneous as donors can be either bilateral or multilateral and differ in size and in their priorities, spatial dependence in aid allocation is tested for four donor groups. Since the degree of spatial dependence might vary with another factor, the spatial lags are finally conditioned on a donor's importance to test whether minor donors are more likely to be spatially dependent on other donors than large donors.

Full sample analysis

Table 19 displays the results of the eligibility stage, estimated by a conditional fixed-effects Logit model. Because this estimation is only based on variation over time, dyads with no change in their allocation status cannot be included in the sample. As a consequence, 659 dyads are lost, 343 because the two members never had a donor-recipient relationship over the whole sample period, and 316 because the donor allocated some positive amount of aid to the recipient in every year the dyad would otherwise be in the sample.[100] Briefly looking at the control variables first, it can be seen that the probability that a certain country receives aid from a given donor is higher if the same donor allocated ODA to this recipient in the last year. As expected, richer countries are less likely to get aid, while the population size of the recipient seems not to influence the probability of receiving aid from a given donor. Nations with a better performance in terms of good governance and countries that receive a higher share of US military grants are more likely to be rewarded with aid. The overall size of the aid budget of a donor increases the probability that a certain recipient receives aid, indicating that bigger donors tend to support a larger number of recipients. There is no evidence that natural disasters increase the probability of receiving aid. A military alliance with the donor seems to reduce the probability to be on a donor's recipients list. This result is at odds with expectation, as there is no theoretical reason for this. First however, a dummy variable is a very crude measure for the military ties between two countries. Second and more importantly, given that the estimation is based on dyad level fixed effects, the estimation of the coefficients is based on variation over time only.

[100] This does not necessarily mean that a donor gives a positive amount of aid to a recipient in all 36 years in the sample (this is true for 64 dyads). For the remaining 252 dyads, the sample period is restricted due to availability the dependent variable or of control variables and there is a positive amount of aid in all years that enter the sample.

Variation in the *Military alliance* variable occurs only if a dyad newly enters or determines such an alliance, which is a rare event: Out of the 1,836 dyads in the sample, only 181 military alliances were forged and 20 terminated, while for 1,636 dyads there is no variation over time.[101] No statistically significant positive effect can be established for bilateral trade.

Turning to the spatial lags, it can be seen that the effects of the spatial lags are in line with expectation and statistically significant at the one percent level throughout if estimated individually. A country is more probable to get aid if other donors give aid to the same recipient (Model I). The same positive effect is found if the other donors are weighted according to their importance (Model II) and if spatial dependence is only allowed between bilateral donors on the one hand and multilateral donors on the other hand (Model III). As expected, all three coefficients become somewhat smaller if all three spatial lags are combined into one model (Model IV). This indicates that even after controlling for a general orientation on other donors, a given donor is more likely to give aid to a certain country if other *important* donors or similar donors give aid to this recipient.

In Model V to VII, the spatial lags with the recipient specific weighting matrices are estimated individually: The coefficients of the spatial lags using the *Bilateral trade* link (Model V) is very small, but still statistically significant, while the effect sizes of the spatial lag using the *Military alliance* link (Model VI) and the spatial lag using the *UN voting similarity* link (Model VII) are considerably larger. In Model VII, all three spatial lags with the recipient specific weighting matrices are combined into one model. Whereas the spatial lag using the bilateral trade link becomes insignificant, the other two remain significant, even though the effect size of the spatial lag using the *Military alliance* link is more than halved. Since general spatial dependence captured by the first set of spatial lags and strategic interaction measured by the spatial lags with the recipient-specific weighting matrices are not mutually exclusive, all spatial lags are estimated together in Model IX. While effect sizes and significance levels of the general spatial lags are hardly affected, only the spatial lag using the *Military alliance* link stays significant in this model.[102] It seems that even after controlling for a general herding behaviour among donors, a given

---

[101] This includes multilateral donors for which the weighted average of military alliances of their members changes over time. One military alliance, namely between the US and Israel, was entered and cancelled during the study period.

[102] Contrary to the Military alliance variable, over time variation in the spatial lag using the *Military alliance* link weighting matrix comes both from changes in the aid allocation of other donors and from changes in the weighting matrix.

donor is still more likely to give aid to a country with which it has a military alliance if other donors also with military ties with the this country decide to allocate some aid to this recipient.

One has to be careful, however, when interpreting these results as causal evidence for spatial contagion. In fact, the estimation procedure automatically controls for all time-invariant factors that influence aid allocation and uses only information on *changes* in the donor-recipient relation, i.e. whether a donor is more likely to start (or stop) giving aid to a recipient if other donors start (or stop) giving aid to this recipient. Furthermore, the model controls for a number of important factors that might trigger a change in such an allocation decision, such as GDP per capita, good governance, or the overall aid budget of a donor. Yet, the list of recipient countries in the sample encompasses only 139 potential candidates and most donors, especially important ones, give aid to a large number of recipients. Given that each donor faces 139 yes-or-no decisions, it could be that several donors simultaneously but independently change the composition of their set of aid partner countries in the same direction.[103] In such a case, the spatial lags would capture spurious effects rather than provide evidence for spatial contagion. These spurious effects are by far less likely to occur in the second stage, where each donor allocates CPA only to those recipients that were selected as partner countries in the first stage. Since the process is not of the yes-or-no decision type, but on the distribution of an aid budget among a (pre-selected) set of recipients, this process entails a much broader range of possible decisions.

[103] In fact, the actual number of decisions included in the sample is lower due to missing information on control variables or the dependent variable.

Table 19: Eligibility stage (first stage) fixed-effects Logit results for all donors

Model	I	II	III	IV	V	VI	VII	VIII	IX
CPA dummy (t-1)	1.354***	1.392***	1.395***	1.337***	1.483***	1.472***	1.382***	1.378***	1.335***
	(44.79)	(46.30)	(46.40)	(44.10)	(50.06)	(49.57)	(45.90)	(45.72)	(44.01)
W: All donors (t-1)	3.669***			2.454***					2.418***
	(23.97)			(10.58)					(6.02)
W: Global aid share (t-1)		1.698***		0.371**					0.356**
		(20.43)		(3.08)					(2.91)
W: Same subgroup of donors (t-1)			1.615***	0.718***					0.720***
			(20.28)	(7.42)					(7.42)
W: Bilateral trade link (t-1)					0.0352***			0.00581	-0.00236
					(4.53)			(0.77)	(-0.32)
W: Military alliance link (t-1)						1.470***		0.575**	0.432*
						(7.56)		(2.96)	(2.19)
W: UN voting similarity link (t-1)							2.848***	2.757***	-0.00197
							(21.67)	(20.27)	(-0.01)
GDP per capita (ln, t-1)	-0.202**	-0.231***	-0.182**	-0.216**	-0.147*	-0.118	-0.188**	-0.180**	-0.209**
	(-2.94)	(-3.36)	(-2.64)	(-3.14)	(-2.13)	(-1.72)	(-2.72)	(-2.61)	(-3.04)
Population (ln, t-1)	0.17	0.18	0.141	0.176	0.159	0.13	0.268	0.292	0.192
	(0.79)	(0.82)	(0.65)	(0.81)	(0.74)	(0.60)	(1.24)	(1.349	(0.88)
Disaster deaths	0.0621	0.0533	0.0368	0.0646	0.0236	0.0154	0.0456	0.0458	0.0632
	(1.22)	(1.03)	(0.70)	(1.25)	(0.47)	(0.31)	(0.90)	(0.91)	(1.22)
Democracy (t-1)	0.0502***	0.0493***	0.0601***	0.0477***	0.0678***	0.0733***	0.0566***	0.0567***	0.0486***
	(5.69)	(5.59)	(6.86)	(5.38)	(7.79)	(8.46)	(6.45)	(6.44)	(5.46)
Bilateral trade (t-1)	-0.295	-0.207	-0.273	-0.265	-0.288*	-0.323*	-0.319*	-0.327*	-0.275
	(-1.91)	(-1.41)	(-1.92)	(-1.72)	(-2.15)	(-2.41)	(-2.15)	(-2.21)	(-1.79)
US military grant share (t-1)	0.0149	0.0216**	0.0261**	0.0148	0.0315***	0.0359***	0.0185*	0.0188*	0.0155
	(1.74)	(2.59)	(3.09)	(1.74)	(3.72)	(4.25)	(2.15)	(2.19)	(1.82)

Table 19: Eligibility stage (first stage) fixed-effects Logit results for all donors (continued)

Model	I	II	III	IV	V	VI	VII	VIII	IX
Diplomatic representation (t-1)	0.275***	0.321***	0.311***	0.265***	0.389***	0.388***	0.296***	0.295***	0.265***
	(3.92)	(4.59)	(4.47)	(3.76)	(5.67)	(5.69)	(4.23)	(4.22)	(3.76)
UN voting similarity (t-1)	0.031	-0.104	0.101	-0.00116	0.142	0.172	0.0445	0.0604	0.0127
	(0.24)	(-0.79)	(0.77)	(-0.01)	(1.10)	(1.32)	(0.34)	(0.46)	(0.10)
Military alliance (t-1)	-0.869***	-1.089***	-1.105***	-0.891***	-1.317***	-1.843***	-0.884***	-1.112***	-1.067***
	(-7.16)	(-9.19)	(-9.43)	(-7.33)	(-11.55)	(-13.56)	(-7.32)	(-7.72)	(-7.26)
Total CPA of donor (ln)	0.952***	0.948***	0.950***	0.955***	0.933***	0.936***	0.953***	0.953***	0.955***
	(30.44)	(30.42)	(30.44)	(30.49)	(30.13)	(30.19)	(30.53)	(30.52)	(30.48)
Observations	45,691	45,691	45,691	45,691	45,691	45,691	45,691	45,691	45,691
Number of dyads	1,836	1,836	1,836	1,836	1,836	1,836	1,836	1,836	1,836

Notes: Dependent variable: dummy taking the value of unity if a donor allocates any positive amount of aid to a recipient; Coefficients displayed; Coefficients on t-1 year dummies not shown; Includes dyad-level fixed effects; Z-values in parenthesis; W: denotes the weighting matrix used for calculating the spatial lags; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

In these second stage estimation models, the dependent variable is recipient *j*'s share of the total CPA commitments of donor *i* in year *t*. Since the estimated model includes a time-lag of the dependent variable, the coefficients displayed are merely the short-run effects. The long-term effects need to take into account the coefficient of the lagged dependent variable. Results are presented in Table 20. Again a brief look at the control variables first: Since the temporal lag of the dependent variable is statistically significant in every model, present aid allocation decisions are not independent of aid given to a beneficiary in the last year. The coefficients indicate that a 10 percentage point increase in last year's share is associated with a 2.2 percentage point increase in the contemporary share. A one percent increase in *GDP pc* reduces the CPA share by 0.56 percentage points (Model I). A change in *Democracy* as well as changes in *Military alliance* is rewarded with a higher share of assistance; however, the size of the first effect is not substantial: a one point increase in the democracy measure (ranging from 2 to 14) leads to a mere 0.06 percentage point higher aid share, whereas the establishment of a military alliance increases the aid share of the respective recipient by more than one percentage point, which equals nearly half of the mean CPA share in the sample (1.92 percentage points). No significant effect is found for *Population*, *Disaster deaths*, *Bilateral trade, Diplomatic representation* and *UN voting similarity*. The effect of the *US military grant share* variable is negative, yet small and not statistically significant in all specifications. As noted by McGillivray and Oczkowski (1992), the lack of an effect of natural disaster fatalities on the share allocated to the affected country could be because aid commitments rather than aid disbursements are used as the dependent variable. Contrary to the latter, which represents the actual amount of funds transferred, aid commitments are future obligations which are less likely to be affected in case of an emergency.

Turning to the spatial lags, there is comprehensive evidence of spatial dependence for the non-recipient specific weighting matrices. As can be seen from Model I, if other donors on average increased the share to a recipient by one percentage point in the previous year, a given donor raises its share by 0.28 percentage points in the short-run and by 0.35 percentage points in the long-run.[104] An increase of one-standard deviation in the spatial lag with which weights all other donors equally (1.19) is associated with a 0.33 percentage point increase in the dependent variable. This represents a 17.2 percent increase from the mean value of the dependent variable (1.92). If instead the impact of other donors is

[104] The long-run effect is calculated as the coefficient of the variable of interest divided by one minus the coefficient of the one-year time lag of the dependent variable (Egger and Merlo 2007).

weighted according to their overall importance in the aid landscape (*W: Global aid share*, Model II), the effect becomes slightly smaller but remains significant at the one percent level. Examining spatial dependence within the groups of bilateral donors and multilateral donors separately, the results show a comparably low effect of 0.08 (Model III). Interestingly, if all three spatial lags measuring herding behaviour among donors are estimated together (Model IV), only the spatial lag which weights other donors according to their importance remains statistically significant. This indicates that donors orientate themselves by the aid allocation decisions of big donors rather than following the example of all other donors equally.

For the spatial lags using recipient-specific weighting matrices, the same pattern can be observed: estimated individually, all three spatial lags are highly significant at least at the one percent level. The effect sizes vary considerably, however. For the spatial lag with the *Bilateral trade* link function, a weighted increase of one percentage point by other donors to a recipient is associated with a 0.04 percentage point increase from an average other donor to the same recipient (Model V). For the spatial lags with the *Military alliance* link und the *UN voting similarity* link, the respective effects are 0.14 percentage points (Model VI) and 0.23 percentage points (Model VII). If all three spatial lags with a recipient-specific weighting matrix are estimated together (Model VIII), the spatial lag with the *Bilateral trade* link loses its significance, whereas the other two are hardly affected.

Finally all six spatial lags are tested simultaneously in Model IX. Only the spatial lags with the *Global aid share* weighting matrix and the *Military alliance* link remain statistically significant and are hardly affected in their coefficient size. A one standard deviation increase in the former variable (1.8 percentage points) has an effect of 0.38 percentage points, while an increase by one standard deviation increase in the latter variable (0.73) is associated with an increase of 0.07 percentage points in the dependent variable. This corroborates the finding from the first stage analysis, namely that spatial dependence in aid allocation pattern is mostly due to an orientation towards major donors rather than due to strategic interaction with other donors. If anything, military considerations matter. Summing up, there is some evidence for positive spatial dependence between donors – and at least no signal for negative spatial dependence which would indicate some kind of donor specialisation on specific countries.

Table 20: Allocation stage (second stage) fixed effects estimation results for all donors

Model	I	II	III	IV	V	VI	VII	VIII	IX
CPA share(t-1)	0.223***	0.221***	0.227***	0.220***	0.227***	0.229***	0.223***	0.223***	0.220***
	(3.57)	(3.52)	(3.69)	(3.49)	(3.71)	(3.77)	(3.59)	(3.59)	(3.49)
W: All donors (t-1)	0.275***			0.0966					0.0785
	(3.68)			(1.18)					(0.40)
W: Global aid share (t-1)		0.251***		0.209*					0.211**
		(3.79)		(2.57)					(2.63)
W: Same subgroup of donors (t-1)			0.0783*	-0.0034					-0.00347
			(2.24)	(-0.09)					(-0.09)
W: Bilateral trade link (t-1)					0.0448**			-0.00491	-0.022
					(2.80)			(-0.32)	(-1.38)
W: Military alliance link (t-1)						0.138**		0.106*	0.0914*
						(2.81)		(2.33)	(2.07)
W: UN voting similarity link (t-1)							0.233***	0.235***	0.0505
							(3.74)	(3.36)	(0.34)
GDP per capita (ln, t-1)	-0.556**	-0.481*	-0.526*	-0.496*	-0.537*	-0.538*	-0.549**	-0.543**	-0.494*
	(-2.62)	(-2.36)	(-2.50)	(-2.42)	(-2.55)	(-2.54)	(-2.59)	(-2.58)	(-2.41)
Population (ln, t-1)	0.0468	-0.0373	0.0518	-0.0282	0.132	0.0867	0.0177	0.017	-0.0612
	(0.12)	(-0.09)	(0.13)	(-0.07)	(0.32)	(0.21)	(0.04)	(0.04)	(-0.16)
Disaster deaths	-0.0233	-0.0447	-0.0489	-0.036	-0.0417	-0.051	-0.0222	-0.0227	-0.0359
	(-0.31)	(-0.59)	(-0.65)	(-0.48)	(-0.55)	(-0.67)	(-0.29)	(-0.30)	(-0.48)
Democracy (t-1)	0.0615***	0.0648***	0.0721***	0.0617***	0.0690***	0.0753***	0.0625***	0.0625***	0.0626***
	(3.35)	(3.57)	(3.98)	(3.40)	(3.67)	(4.15)	(3.43)	(3.36)	(3.39)
Bilateral trade (t-1)	-0.604	-0.635	-0.672	-0.618	-0.643	-0.672	-0.613	-0.617	-0.624
	(-1.28)	(-1.35)	(-1.44)	(-1.31)	(-1.38)	(-1.45)	(-1.30)	(-1.31)	(-1.32)
US military grant share (t-1)	-0.0561*	-0.0416	-0.0575*	-0.0438	-0.0588*	-0.0582*	-0.0556*	-0.0564*	-0.0434
	(-2.02)	(-1.55)	(-2.10)	(-1.67)	(-2.14)	(-2.13)	(-2.00)	(-2.02)	(-1.66)

Table 20: Allocation stage (second stage) fixed effects estimation results for all donors (continued)

Model	I	II	III	IV	V	VI	VII	VIII	IX
Diplomatic representation (t-1)	0.221	0.226	0.231	0.222	0.23	0.238	0.227	0.228	0.224
	(1.57)	(1.62)	(1.65)	(1.58)	(1.64)	(1.70)	(1.61)	(1.62)	(1.61)
UN voting similarity (t-1)	0.369	0.386	0.366	0.382	0.375	0.389	0.371	0.383	0.392
	(1.42)	(1.48)	(1.40)	(1.47)	(1.43)	(1.49)	(1.42)	(1.47)	(1.50)
Military alliance (t-1)	1.141***	1.111***	1.157***	1.110***	1.174***	1.103***	1.146***	1.088***	1.057***
	(6.47)	(6.28)	(6.50)	(6.28)	(6.63)	(6.04)	(6.49)	(6.02)	(5.82)
Constant	5.431	6.112	5.316	6.053	4.11	4.886	5.858	5.817	6.552
	(0.81)	(0.92)	(0.79)	(0.91)	(0.61)	(0.72)	(0.87)	(0.88)	(1.00)
Observations	30,396	30,396	30,396	30,396	30,396	30,396	30,396	30,396	30,396
Number of dyads	2,152	2,152	2,152	2,152	2,152	2,152	2,152	2,152	2,152
R-squared	0.089	0.090	0.087	0.090	0.087	0.087	0.089	0.089	0.091

Notes: Dependent variable: Recipient *j*'s share at donor *i*'s total CPA in year *t*; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on (t-1) year dummies not displayed; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

Results of a donor-group specific analysis

So far, only average values for all donors have been analysed. Such a pooled analysis might veil important results if donors react differently to aid allocation decisions of other donors. This is particularly relevant as there are theoretical arguments both for negative and positive spatial dependence among donors. No clear-cut results in an aggregate estimation might be due to the possibility that some donors react with less aid to a recipient if other donors allocate more, while other donors give a higher share of their aid to this recipient. Such varying dependencies could neutralise if all donors are estimated together and only net-effects remain visible. Not only can bilateral and multilateral donors be distinguished, but donors also differ with respect to their aid allocation agenda. Apart from the traditional big Western donors, especially the group of like-minded countries deserves closer attention as these countries enjoy the reputation of being less focussed on donor interest and of giving aid more strongly according to recipient needs. In the following, the extent of spatial dependence is scrutinised for donor groups. In order to allow all coefficients to differ for each donor group, the estimation is run for each subsample separately rather than interacting the spatial lags with donor group specific dummies.

Table 21 shows the effects of different spatial lags for bilateral donors. Due to space restrictions, the complete estimation results are not shown. The set of control variables is as in the estimations shown in Table 20. The overall pattern of this subsample is the same as for the whole sample – estimated individually, with the exception of the spatial lag with the same subgroup weighting matrix, all other spatial lags are statistically significant at least at the five percent level. If all three spatial lags are estimated simultaneously (Model IV), only the spatial lag in which ODA of other donors is weighted according to their relative importance remains significant. This spatial lag is still significant if all six spatial lags are estimated together in Model IX. The only other variable with a significant effect is the spatial lag with the *Military alliance* link. The effect size, however, is only one fourth of the spatial lag with the *Global aid share* weighting matrix. Turning to the results for multilateral organisations (Table 22), it can be seen that aid allocation decisions of these donors are time-dependent, but there is no evidence that multilateral donors are influenced by the aid allocation decisions of neither other multilaterals nor bilateral donors. Since estimations are based on fewer observations the detection of a statistical significant effect is less likely. The striking result is that the spatial lags fail to reach conventional significance levels even if estimated individually. This indicates that multilateral organisations are fully independent in their aid allocation decisions.

Table 21: Allocation stage (second stage) fixed effects estimation results for bilateral donors

Model	I	II	III	IV	V	VI	VII	VIII	IX
CPA share(t-1)	0.226**	0.224**	0.229**	0.223**	0.230***	0.232***	0.226**	0.226**	0.223**
	(3.19)	(3.16)	(3.28)	(3.14)	(3.32)	(3.37)	(3.21)	(3.21)	(3.13)
W: All donors (t-1)	0.267**			0.0949					0.0386
	(3.23)			(1.06)					(0.17)
W: Global aid share (t-1)		0.251***		0.218*					0.223*
		(3.46)		(2.41)					(2.50)
W: Same subgroup of donors (t-1)			0.0835	-0.0174					-0.0174
			(1.84)	(-0.36)					(-0.35)
W: Bilateral trade link (t-1)					0.0436*			-0.00645	-0.0227
					(2.47)			(-0.37)	(-1.26)
W: Military alliance link (t-1)						0.0993***		0.0676*	0.0550*
						(3.96)		(2.51)	(2.07)
W: UN voting similarity link (t-1)							0.230**	0.237**	0.0862
							(3.29)	(2.98)	(0.48)
Observations	25,831	25,831	25,831	25,831	25,831	25,831	25,831	25,831	25,831
Number of dyads	1,859	1,859	1,859	1,859	1,859	1,859	1,859	1,859	1,859
R-squared	0.087	0.089	0.085	0.089	0.085	0.085	0.087	0.087	0.089

Notes: To save space, results for control variables (as in Table 20) are not shown; Dependent variable: Recipient *j*'s share at donor *i*'s total CPA in year *t*; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on (t-1) year dummies not displayed; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

Table 22: Allocation stage (second stage) fixed effects estimation results for multilateral donors

Model	I	II	III	IV	V	VI	VII	VIII	IX
CPA share(t-1)	0.191***	0.188***	0.198***	0.188***	0.195***	0.195***	0.193***	0.191***	0.187***
	(3.74)	(3.73)	(3.80)	(3.80)	(3.74)	(3.74)	(3.75)	(3.71)	(3.78)
W: All donors (t-1)	0.316			0.176					0.353
	(1.86)			(1.26)					(1.35)
W: Global aid share (t-1)		0.232		0.146					0.118
		(1.73)		(1.25)					(1.06)
W: Same subgroup of donors (t-1)			0.0516	0.0162					0.0139
			(1.31)	(0.34)					(0.29)
W: Bilateral trade link (t-1)					0.0506			0.00522	-0.0165
					(1.51)			(0.20)	(-0.54)
W: Military alliance link (t-1)						0.294		0.262	0.252
						(1.83)		(1.61)	(1.55)
W: UN voting similarity link (t-1)							0.235	0.205	-0.126
							(1.80)	(1.57)	(-0.79)
Observations	4,565	4,565	4,565	4,565	4,565	4,565	4,565	4,565	4,565
Number of dyads	293	293	293	293	293	293	293	293	293
R-squared	0.142	0.142	0.138	0.143	0.139	0.14	0.141	0.143	0.145

Notes: To save space, results for control variables (as in Table 20) are not shown; Dependent variable: Recipient *j*'s share at donor *i*'s total CPA in year *t*; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on (t-1) year dummies not displayed; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

The group of bilateral donors can be further split up into big Western donors, such as France, Germany, the United Kingdom and the United States, and into the group of like-minded countries, which encompasses the Scandinavian countries, the Netherlands and Canada, whereas other bilateral donors are smaller European donors, Australia and New Zealand as well as the Republic of Korea. The results for all three groups are shown in Tables 23 to 25. The spatial lags with the global spatial lags are individually statistically significant for the first two subgroups, but the effects are generally larger for the group of like-minded countries. If these three spatial lags are combined into one Model (IV), some differences are revealed. While for the big Western donors the spatial lag with the importance-weighted ODA of other donors remains significant, for like-minded donors, the spatial lag with the *Same subgroup* weighting matrix survives. Since the big Western donors are also the most important players in the global aid landscape, the spatial lag with the *Global aid share* weighting matrix is quite similar to the weighting matrix with the *Same subgroup* weighting matrix for these donors. Interestingly, like-minded donors seem not to follow the example of the big Western donors but rather react to changes in aid allocation by other like-minded donors. Regarding the recipient specific-weighting matrices, another difference becomes apparent: while the spatial lag with the *Military alliance* link is significant for the group of big Western donors, it has no effect for like-minded donors (Model VI). If all six spatial lags are incorporated into one Model (IX), the spatial lags with the *Global aid share* weighting matrix and with the *Military alliance* link weighting matrix remain significant in the Big Western donors sample, whereas for like-minded countries, this is true for the spatial lag that weights all other donors equally and for the spatial lag with the *Same subgroup* weighting matrix. Finally, for the group of other bilateral donors, there is no evidence for herding behaviour as all three spatial lags with global weighting matrices are insignificant. Only the spatial lag with the Military alliance link weighting matrix is significant if estimated individually (Model VI) or together with the other two spatial lags with a recipient-specific weighting matrix (Model VIII). The effect however does not remain significant in the most comprehensive model IX. This shows that donors of this group pay little attention to aid allocation decisions of other donors, but there is some evidence that they possibly use aid strategically to secure their influence in recipients with whom they have a military alliance.

Taken together, the results provide considerable evidence for herding behaviour in aid allocation, but little evidence for strategic interaction among donors when giving aid. The general pattern is that donors follow the example of other large donors when allocating their aid to different recipient countries. However, this result is driven by big Western donors as there is no spatial dependence for multilateral organisations and like-minded donors' allocation decision seem to be only geared to ODA commitments of other donors of the same group. If anything, strategic interaction among donors takes place with regard to military strategic targets: a donor allocates a larger share of its aid to a recipient with which it has a signed a military alliance if another donor with such an alliance gives more aid to this recipient.

Table 23: Allocation stage (second stage) fixed effects estimation results for big Western donors

Model	I	II	III	IV	V	VI	VII	VIII	IX
CPA share(t-1)	0.125*	0.128*	0.134*	0.124*	0.137*	0.140*	0.127*	0.127*	0.124*
	(2.06)	(2.11)	(2.25)	(2.04)	(2.27)	(2.33)	(2.08)	(2.08)	(2.02)
W: All donors (t-1)	0.332***			0.23					0.13
	(3.37)			(1.82)					(0.33)
W: Global aid share (t-1)		0.240**		0.193*					0.194*
		(3.16)		(1.97)					(1.98)
W: Same subgroup of donors (t-1)			0.101*	-0.0689					-0.0632
			(2.47)	(-1.31)					(-1.21)
W: Bilateral trade link (t-1)					0.0490*			-0.0117	-0.0241
					(2.11)			(-0.49)	(-0.83)
W: Military alliance link (t-1)						0.113***		0.0809**	0.0736**
						(4.42)		(2.76)	(2.56)
W: UN voting similarity link (t-1)							0.289***	0.300**	0.124
							(3.65)	(3.30)	(0.51)
Observations	10,256	10,256	10,256	10,256	10,256	10,256	10,256	10,256	10,256
Number of dyads	576	576	576	576	576	576	576	576	576
R-squared	0.06	0.06	0.055	0.061	0.055	0.054	0.059	0.06	0.062

Notes: To save space, results for control variables (as in Table 20) are not shown; Dependent variable: Recipient *j*'s share at donor *i*'s total CPA in year *t*; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on (t-1) year dummies not displayed; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

Table 24: Allocation stage (second stage) fixed effects estimation results for like-minded donors

Model	I	II	III	IV	V	VI	VII	VIII	IX
CPA share(t-1)	0.119***	0.124***	0.105***	0.0996***	0.135***	0.140***	0.122***	0.122***	0.0992***
	(4.51)	(4.49)	(4.34)	(4.11)	(4.94)	(5.00)	(4.57)	(4.55)	(4.11)
W: All donors (t-1)	0.418***			0.157					0.437*
	(5.50)			(1.82)					(2.33)
W: Global aid share (t-1)		0.252***		0.0662					0.0638
		(4.47)		(1.04)					(1.00)
W: Same subgroup of donors (t-1)			0.273***	0.221***					0.216***
			(5.49)	(3.86)					(3.68)
W: Bilateral trade link (t-1)					0.0629**			-0.0213	-0.014
					(2.76)			(-0.91)	(-0.57)
W: Military alliance link (t-1)						0.00695		-0.0966	-0.0234
						(0.13)		(-1.31)	(-0.35)
W: UN voting similarity link (t-1)							0.330***	0.365***	-0.226
							(5.09)	(4.61)	(-1.59)
Observations	7,836	7,836	7,836	7,836	7,836	7,836	7,836	7,836	7,836
Number of dyads	464	464	464	464	464	464	464	464	464
R-squared	0.079	0.074	0.088	0.091	0.068	0.065	0.076	0.076	0.091

Notes: To save space, results for control variables (as in Table 20) are not shown; Dependent variable: Recipient *j*'s share at donor *i*'s total CPA in year *t*; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on (t-1) year dummies not displayed; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

Table 25: Allocation stage (second stage) fixed effects estimation results for other bilateral donors

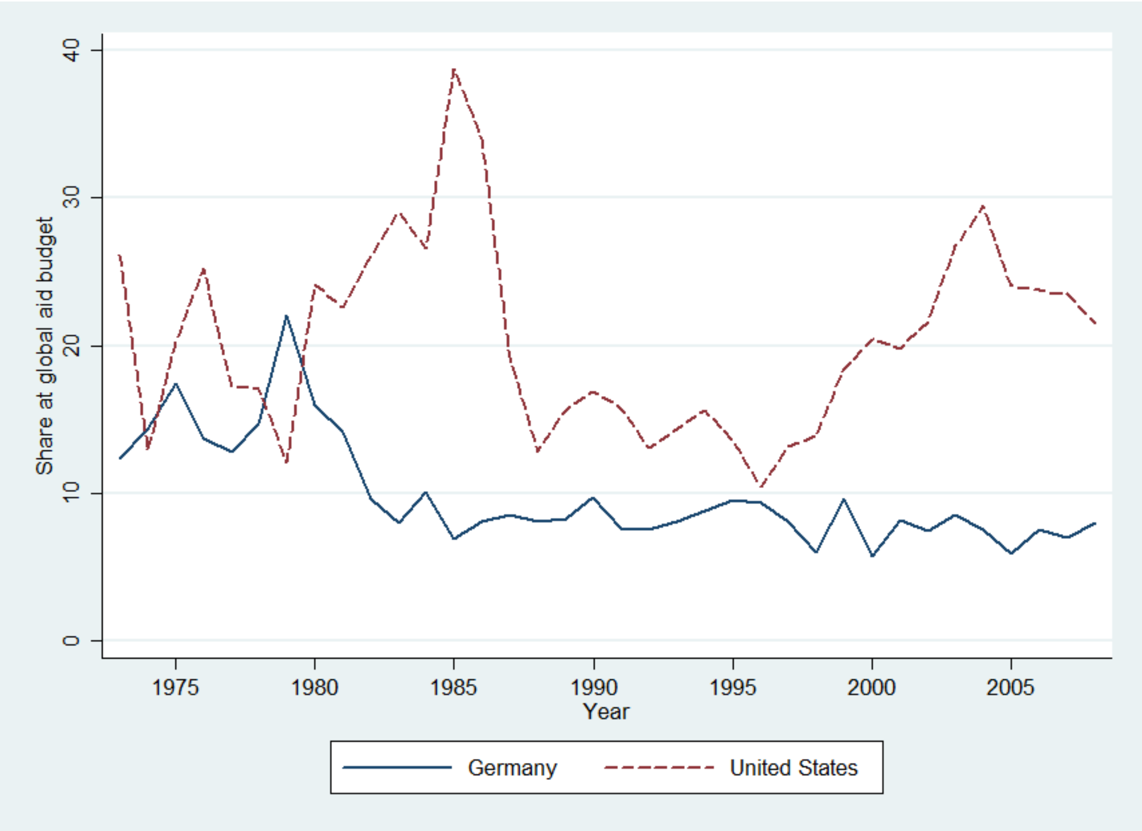
Model	I	II	III	IV	V	VI	VII	VIII	IX
CPA share (t-1)	0.266**	0.262*	0.267**	0.263*	0.266**	0.267**	0.265**	0.265**	0.262**
	(2.64)	(2.56)	(2.68)	(2.58)	(2.66)	(2.68)	(2.64)	(2.64)	(2.58)
W: All donors (t-1)	0.121			-0.297					-0.52
	(0.61)			(-0.82)					(-0.92)
W: Global aid share (t-1)		0.318		0.44					0.439
		(1.85)		(1.84)					(1.86)
W: Same subgroup of donors (t-1)			-0.0342	0.0197					0.0204
			(-0.28)	(0.13)					(0.12)
W: Bilateral trade link (t-1)					0.0321			0.0139	0.0048
					(0.90)			(0.30)	(0.10)
W: Military alliance link (t-1)						0.697***		0.587*	0.377
						(3.43)		(2.15)	(1.29)
W: UN voting similarity link (t-1)							0.119	0.087	0.196
							(0.63)	(0.36)	(0.38)
Observations	7,739	7,739	7,739	7,739	7,739	7,739	7,739	7,739	7,739
Number of dyads	819	819	819	819	819	819	819	819	819
R-squared	0.153	0.155	0.153	0.156	0.153	0.153	0.153	0.153	0.156

Notes: To save space, results for control variables (as in Table 20) are not shown; Dependent variable: Recipient *j*'s share at donor *i*'s total CPA in year *t*; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on (t-1) year dummies not displayed; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

Conditional effects of spatial dependence

The donor-group specific analysis provided a first indication that the degree of spatial dependence is not the same for all donors; however, it does not show whether the effect varies systematically across individual donors. One argument for positive spatial dependence is the reduction of uncertainty necessarily involved in aid projects. Since larger donors implement a higher number of aid projects and can employ a larger administrative workforce to assess the potential effectiveness of projects, they could face a lower risk of not reaching the aspired project goals than smaller donors. The same is true if these donors are able to influence the recipient's behaviour due to their financial clout more strongly. Therefore, donors with a smaller absolute aid budget might follow the example of larger donors but to a lesser extent vice versa. The positive and significant spatial lag *Global aid share* in the previous subsections already points into that direction, but it does not show whether spatial dependence systematically differs with a donor's importance. As a measure of importance, the share of donor *i* at the total global CPA budget in year *t* is taken. This measure is time-variant to reflect the changing importance of a donor over time. As can be seen form Figure 23, the relative importance of a donor fluctuates considerably over time. Germany's share peaked in 1979 with 22 percent of the total global aid commitment in that year, but decreased to 8 percent in 1983 and remained at that level since then. The relative importance of the United States reached its height in 1985 with nearly 39 percent followed by a descent to 12.8 percent in 1988 and a new increase to almost 30 percent in 2004.

Figure 23: Share of a donor's aid at the total global aid budget



All spatial lags are interacted with this measure for the importance of a donor. As outlined above, at least for the three spatial lags measuring herding behaviour, a negative coefficient for the interaction effect is anticipated. For the remaining three spatial lags with the recipient-specific weighting matrices, which capture strategic interaction among donors, the effect is not clear. If one is willed to assume that larger donors use their aid more strongly to pursue their strategic goals, a positive interaction effect would be expected. The results of the conditional spatial lags are presented in Table 26. The basic effect of the variable *Aid share donor* is negative and significant in seven out of nine estimation models, indicating that on average larger donors give a lower share of their aid to a single recipient. This is in line with the expectation that smaller donors tend to focus their assistance on a smaller subset of countries, thereby giving a higher average share to each recipient. As in previous estimations, the non-interacted effects are positive and statistically significant at least at the five percent level and the inclusion of the interaction effects inflates the coefficients of the non-interacted spatial lags slightly. Looking at the three spatial lags that measure herding behaviour (Model I to III), it can be seen that all interaction effects are negative as expected, but only statistically significantly for the spatial lag using the *Global aid share* and the *Same subgroup* weighting matrix. This provides evidence that smaller donors gear their aid allocation decisions to larger donors, but not the other way round: the larger a donor, the less strongly it is influenced by others. Turning to the spatial lag measuring strategic interaction (Model IV to VI), the same pattern can be observed. Again, smaller donors are more spatially dependent on others than large donors. While this is at odds with the possible theoretical argument provided above, one has to bear in mind that the interaction variable measures a donor's importance in the global aid landscape rather than the importance of aid for a given donor. This would be measured by aid as a share of GDP and will be analysed further below. If all spatial lags and their interaction effects are estimated simultaneously, only the spatial lag with the *Global aid share* weighting matrix and its interaction effect remain significant.

Table 26: Conditional spatial lag effects (interacted with aid share donor, allocation stage fixed effects estimation results for all donors)

Model	I	II	III	IV	V	VI	VII	VIII	IX
CPA share(t-1)	0.223***	0.220***	0.226***	0.219***	0.227***	0.229***	0.223***	0.223***	0.219***
	(3.57)	(3.51)	(3.68)	(3.47)	(3.71)	(3.77)	(3.59)	(3.59)	(3.47)
W: All donors (t-1)	0.325***			-0.0347					-0.131
	(3.71)			(-0.29)					(-0.47)
W: All donors (t-1) x Aid share donor	-0.00903			0.0213*					0.0326
	(-1.73)			(2.28)					(1.80)
W: Global aid share (t-1)		0.344***		0.346**					0.349**
		(4.01)		(2.81)					(2.86)
W: Global aid share (t-1) x Aid share donor		-0.0169**		-0.0253*					-0.0256*
		(-2.97)		(-2.37)					(-2.39)
W: Same subgroup of donors (t-1)			0.125**	0.038					0.0386
			(2.77)	(0.72)					(0.73)
W: Same subgroup of donors (t-1) x Aid share donor			-0.00806**	-0.00421					-0.00421
			(-2.79)	(-0.99)					(-0.98)
W: Bilateral trade link (t-1)					0.0622**			0.00654	-0.0139
					(2.85)			(0.28)	(-0.63)
W: Bilateral trade link (t-1) x Aid share donor					-0.00318*			-0.00216	-0.000831
					(-2.07)			(-1.32)	(-0.58)
W: Military alliance link (t-1)						0.165*		0.117	0.0719
						(2.09)		(1.48)	(0.91)
W: Military alliance link (t-1) x Aid share donor						-0.00198		-0.00032	0.00219
						(-0.80)		(-0.12)	(0.80)
W: UN voting similarity link (t-1)							0.277***	0.264**	0.112
							(3.60)	(2.99)	(0.51)
W: UN voting similarity link (t-1) x Aid share donor							-0.00774	-0.00506	-0.00983
							(-1.61)	(-0.95)	(-0.72)

Table 26: Conditional fixed effects (continued)

Model	I	II	III	IV	V	VI	VII	VIII	IX
Aid share donor	-0.0189*	-0.0079	-0.0188*	-0.014	-0.0215**	-0.0278***	-0.0200*	-0.0182*	-0.0131
	(-2.12)	(-0.82)	(-2.41)	(-1.58)	(-2.81)	(-3.88)	(-2.29)	(-2.09)	(-1.48)
Observations	30,396	30,396	30,396	30,396	30,396	30,396	30,396	30,396	30,396
Number of dyads	2,152	2,152	2,152	2,152	2,152	2,152	2,152	2,152	2,152
R-squared	0.089	0.092	0.088	0.093	0.088	0.087	0.089	0.090	0.093

Notes: To save space, results for control variables (as in Table 20) are not shown; Dependent variable: Recipient *j*'s share at donor *i*'s total CPA in year *t*; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on (t-1) year dummies not displayed; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

Discussion of results and contrasting them with previous studies

As mentioned above, all previous empirical studies which examined spatial dependence did so using a unitary weighting matrix. This is the first analysis that not only used more elaborated global weighting matrices, but also introduced recipient-specific weighting matrices to test for strategic interaction among donor countries and multilateral organisations. As a consequence, only the results of the spatial lag using the *All donors* weighting matrix can be contrasted with the findings of previous studies. As shown in the literature overview in Table 17, most studies find at least some evidence for spatial dependence. While in the analysis for British aid allocation by McGillivray and Oczkowski (1992) spatial dependence was found only in the allocation stage but not the eligibility stage, in the present analysis the spatial lag with the *All donors* weighting matrix is positive and highly significant in both stages if only British aid commitments are regarded. However, this difference could be easily explained by the longer time series used in our sample and the larger set of recipient countries.[105] Tarp et al. (1998) find that Danish aid spatially depends on other donors in both stages. This result is confirmed in the new dataset used here. Vázquez (2008) examines the determinants of Spanish aid allocation decisions and concludes that the selection of recipient countries is affected by decisions of other countries, but that the aid allocation is independent. Again, both results can

[105] In fact, if the sample is restricted to 1980 to 1987, the effect in the first stage becomes insignificant while the effect in the second stage remains significant.

be reproduced. The last donor-specific study is by Maurini and Settimo (2009) for Italy who find positive spatial dependence in both stages. While the positive effect in the first stage can be detected in the first stage, the spatially lagged dependent variable is not significant in the second stage. However, the analysis of Maurini and Settimo uses a Tobit model whereas the empirical approach applied here is a fixed-effects OLS. If both stages are estimated simultaneously by a random-effects Tobit model, the spatial lag using the *All donors* weighting matrix becomes positive and statistically significant at the one percent level. Yet, as argued above, including dyad level fixed effects is crucial when analysing spatial dependence.

Turning to the studies that also use a larger dataset with various donors and recipients, Berthélemy and Tichit (2004) also find positive spatial dependence in their full model, but negative spatial dependence during the 1980s. The latter result cannot be replicated, as the spatial lag is positive and significant even with a Tobit model. Also, using a Tobit model, no negative spatial dependence is found neither for the group of Belgium, Ireland and Italy together nor for any donor individually. This might be explained by differences in the definition of both the dependent variable and the measure of spatial dependence. While Berthélemy and Tichit (2004) analyse the determinants of aid per capita, here the share of aid to a recipient at the total aid budget of a donor is used. Furthermore, Berthélemy and Tichit include aggregate ODA per capita by other donors to capture other donors' behaviour. Hence, they implicitly use a non row-standardised weighting matrix, whereas here the average aid share of other donors is used. Finally, Berthélemy and Tichit use all ODA commitments, but for this analysis, these amounts are adjusted by commitments related to emergency and food aid as well as to debt relief. Berthélemy (2006) finds a negative spatial dependence for aid from other bilateral donors in a two-part model with fixed effects. Since his sample only covers bilateral donors, this is conceptionally equivalent to the spatial lag with the *Same subgroup* weighting matrix for a sample of bilateral donors only, which is presented in Model III of Table 21 above. This spatial lag, however, is not negative, but positive yet insignificant. The same is true if the analysis is restricted to the group of donors for which Berthélemy finds negative spatial dependence (Australia, France, Italy, Japan, UK and USA). Potential reasons are the same as for the previous study by Berthélemy and Tichit (2004). Finally, Claessens et al. (2009) find evidence for positive spatial dependence in their dynamic panel analysis estimated via GMM, which is in line with the findings provided above. Taken together, the results obtained in this study generally confirm previous studies, but also provide a considerable extension.

3.6. Conclusion

The aim of this paper is to analyse how aid allocation by one donor is influenced by the aid distribution of other donors. There are several reasons why this might be the case: Positive spatial dependence can occur if a donor uses foreign aid allocation decisions by other donors as a guiding line for its own decision to reduce uncertainty necessarily involved in aid projects. Another explanation for positive spatial dependence can be provided by strategic interaction among donors, since ODA is at least partly given for strategic reasons. If competing donors use their aid to secure national interests, every donor should respond to the behaviour of its peers by adapting its own aid allocation to the allocation decisions of other donors. Negative spatial dependence, i.e. if one donor reduces the relative importance of a given recipient after other donors allocated a larger share of their aid budget to this recipient, can indicate a certain degree of donor specialisation over time.

The decision process is modelled as a two-step procedure, in which a donor in the first step compiles the recipient list, i.e. the set of countries that will receive some positive amount of aid. In a second step, the aid budget is distributed to these countries. Spatial dependence is analysed in both stages. In the first stage, there is comprehensive evidence that the probability that a donor gives some positive amount to a certain recipient is higher if this recipient also receives aid from other donors. This result holds true after controlling for various aspects of recipient need. A positive and significant effect can be established both for the spatial lags using weighting matrices which measure herding behaviour as well as for the spatial lags using recipient-specific weighting matrices. The latter set of weighting matrices models strategic interaction between donors with respect to economic, military-strategic and political considerations.

The estimation results for the second stage suggest that the aid allocation of other donors matters in fact: if other donors increase (decrease) the relative aid budget for a recipient, a given donor is also likely to allocate relatively more (less) to the same recipient. A disaggregated analysis for donor groups shows that there is no evidence that multilateral organisations spatially depend on other donors. While the big Western donors seem to gear their aid allocation decisions to aid allocation of other important donors, like-minded donors are mainly influenced by other like-minded donors. Generally, the degree of spatial dependence decreases with the own importance, i.e. small donors are more responsive to past changes in aid allocation of others than large donors. Negative spatial dependence would indicate that aid flows by one donor are a substitute rather than a complement to foreign aid by other donors. No evidence, however, is found for such negative spatial dependence in any model or specification. While this clearly shows that there is no donor specialisation, one has to be careful to interpret this as evidence for a lack of donor coordination, since it does not provide insights in how aid flows by various donors to one recipient are dovetailed on a project-level.

Yet, while comprehensive evidence for spatial dependence is found, it points into the direction of simple herding behaviour rather than strategic interaction among donors. At least for economic and political interest, the results suggest that a donor does not allocate more aid to an important trade partner (or to a country with a similar voting pattern in the UN general assembly), if another donor, that also has close trade-ties with (or a similar voting behaviour as) the same recipient, increases its aid share to this recipient. If anything, only for a military alliance between a donor and a recipient tentative evidence for such behaviour can be detected. However, the effect size for the spatial lag is less than half of the size for the herding measure. Summarising the evidence, it is clear that a donors aid giving is in fact influenced by others donors. Yet, donors do not pay the same attention to the actions of any other donor. The results clearly demonstrate that a donor’s behaviour is geared to the behaviour of important players in the aid game: Donors seem in fact to be sheep – following the bellwethers.

References

Agénor, Pierre-Richard and Joshua Aizenman (2007), *Aid Volatility and Poverty Traps*, NBER Working Paper Series, 13400, Cambridge, MA: National Bureau of Economic Research.

Aldasoro, Inaki, Peter Nunnenkamp and Rainer Thiele (2010), *Less Aid Proliferation and More Donor Coordination? The Wide Gap Between Words and Deeds*, Journal of International Development, 22, pp. 920-940.

Alesina, Alberto and David Dollar (2000), Who Gives Foreign Aid to Whom and Why?, *Journal of Economic Growth*, 5(1), pp. 33-63.

Allison, Graham (1971), *Essence of Decision*, Boston, MA: Little, Brown.

Anselin, Luc (1988), *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer.

Arellano, Christina, Aleš Bulíř, Timothy Labne, and Leslie Lipschitz (2009), The dynamic implications of foreign aid and its variability, *Journal of Development Economics*, 88, pp. 87-102.

Arndt, Channing (2000), Technical Cooperation, in: Tarp, Finn (ed), *Foreign Aid and Development: Lessons Learnt and Directions for the Future*, London: Routledge, pp. 120-138.

Berg, E. (1997), Dilemmas in Donor Aid Strategies, in: Gwin, C. and Nelson, J.M. (eds), *Perspectives on Aid and Development*, Baltimore: John Hopkins University Press, pp. 79-94.

Berthélemy, Jean-Claude (2006), Bilateral Donors' Interest vs. Recipients' Development Motives in Aid Allocation: Do All Donors Behave the Same?, *Review of Development Economics*, 10(2), pp. 179-194.

Berthélemy, Jean-Claude and Ariane Tichit (2004), Bilateral donors' aid allocation decisions – a three-dimensional panel analysis, *International Review of Economics and Finance*, 13, pp. 253-274.

Bigsten, Arne (2006), *Donor coordination and the uses of aid*, Göteborg: Göteborg University.

Bulíř, Aleš and A. Javier Hamann (2003), Aid volatility: An empirical assessment, *IMF Staff Papers*, 50(1), pp. 64-89.

Bulíř, Aleš and A. Javier Hamann (2008), Volatility of Development Aid: From the Frying Pan into the Fire?, *World Development*, 36(10), pp. 2048-2066.

Cassen, Robert (1994), *Does Aid Work?,* Oxford: Oxford University Press.

Claessens, Stijn, Danny Cassimon, and Bjorn Van Campenhout (2009), Evidence on Changes in Aid Allocation Criteria, *The World Bank Economic Review*, 23(2), pp. 185-208.

Djankov, Simeon, Jose G. Montalvo, and Marta Reynal-Querol (2009), Aid with multiple personalities, J*ournal of Comparative Economics*, 37, pp. 217-229.

Dreher, Axel, Peter Nunnenkamp and Rainer Thiele (2007), Does US aid buy UN general assembly votes? A disaggregated analysis, *Public Choice*, 136(1-2), pp. 139-164.

Dudley, Leonard and Claude Montmarquette (1976), A Model of the Supply of Bilateral Foreign Aid, *American Economic Review*, 66(1), pp. 132-142.

Edwards, Sebastian and Sweder van Wijnbergen (1989), Disequilibrium and structural adjustment, in: Schulz, T. and John Strauss (eds), *Handbook of Development Economics*, Vol. 2, Cambridge: North-Holland, pp. 1481-1533.

Egger, Peter and Valeria Merlo (2007), The Impact of Bilateral Investment Treaties on FDI Dynamics, *The World Economy*, 10, pp. 1536-1549.

EM-DAT (2010), *EM-DAT: The International Disaster Database*, Brussels: Centre for Research on the Epidemiology of Disasters.

European Commission (2009), *EU Toolkit for the implementation of complementarity and division of labour in development policy*, Brussels: European Commission Directorate-General for Development and Relations with African, Caribbean and Pacific States.

Fatás, Antonio and Ilian Mihov (2008), Fiscla Discipline, Volatility and Growth, in: Perry, G., L. Serven, and R. Suescun (eds), *Fiscal Policy, Stabilization and Growth*, Washington, DC: World Bank.

Fielding, David and George Mavrotas (2005), *The Volatility of Aid*, WIDER Discussion Paper No. 2005/06, Helsinki: UNU-WIDER.

Franzese, Robert and Jude C. Hays (2008), Spatial-Econometric Models of Interdependence. Book prospectus.

Freedom House (2009), Annual Surveys of Freedom Country Ratings 1972 to 2009, New York: Freedom House.

.

Frot, Emmanuel and Javier Santiso (2011), Herding in Aid Allocation, *Kyklos*, 64(1), pp.54-74.

Galton, Francis (1889), Discussion on Edward B. Tylor's 'On a Method of Investigating the Development of Institutions, Applied to Laws of Marriage and Descent', *Journal of the Anthropological Institute of Great Britain and Ireland*, 18, pp. 270–272.

Gartzke, Erik (2010), The Affinity of Nations: Similarity of State Voting Positions in the UNGA, available at: http://dss.ucsd.edu/~egartzke/htmlpages/data.html, last accessed: 02.07.2010.

Hopkins, Raymond (2000), Political Economy of Foreign Aid, in: Tarp, F. (ed.) *Foreign Aid and Development: Lessons Learnt and Directions for the Future*, London: Routledge, pp. 423-449.

Katz, Ethan (2001), Bias in Conditional and Unconditional Fixed Effects Logit Estimation, *Political Analysis*, 9(4), pp. 379-384.

Kharas, Homi (2008), Measuring the Cost of Aid Volatility, *Wolfenson Center for Development Working Paper*, No. 3, Washington, DC: Brookings.

Knack, Stephen and Aminur Rahman (2008), Donor Fragmentation, in: Easterly, William (ed), *Reinventing Foreign Aid*, Cambridge, MA: The MIT Press, pp. 333-347.

Leeds, Brett Ashley, Jeffrey M. Ritter, Sara McLaughlin Mitchell, and Andrew G. Long (2002), Alliance Treaty Obligations and Provisions, 1815-1944, *International Interactions*, 28, pp. 237-260.

Lensink, Robert and Oliver Morrisey (2000), Aid instability as a measure of uncertainty and the positive impact of aid on growth, *Journal of Development Studies*, 36, pp. 31-49.

Lensink, Robert and Howard White (2001), Are There Negative Returns to Aid?, *Journal of Development Studies*, 37(6), pp. 42-64.

Maizels, Alfred and Machiko Nissanke (1984), Motivations for Aid to Developing Countries, *World Development*, 12(9), pp. 879-900.

Manning, W.G., N. Duan, and W.H. Rogers (1987), Monte-Carlo Evidence on the Choice Between Sample Selection and Two-Part Models, *Journal of Econometrics*, 385, pp. 59-82.

Maurini, Claudia and Riccardo Settimo (2009), Assessing the allocation of Italian Foreign Aid, *Questioni di Economia e Finanza*, 43, Bank of Italy.

McGillivray, Mark and Edward Oczkowski (1992), A two-part sample selection model of British bilateral foreign aid allocation, *Applied Economics*, 24, pp. 1311-1319.

McGillivray, Mark and White, Howard (1993), Explanatory Studies of Aid Allocation Among Developing Countries: A Critical Survey, *Working Paper Series No 148*, mimeo.

McKinlay, R.D. and R. Little (1977), A Foreign Policy Model of U.S. Bilateral Aid Allocation, *World Politics*, 30(1), pp. 58-86.

Neumayer, Eric (2003), *The Pattern of Aid Giving: The impact of good governance on development assistance*, London: Routledge.

Neumayer, Eric and Thomas Plümper (2010), Spatial Effects in Dyadic Data, *International Organization*, 64(1), pp. 145-166.

Nickell, Stephen (1981), Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49, pp. 1417-1426.

Nielson, Daniel L., Ryan M. Powers, and Michael J. Tierney (2010), *Broad Trends in Foreign Aid: Insights from PLAID 1.9.1.*, Conference paper presented at Aid Transparency and Development Finance: Lessons and Insights from AidData, Oxford, U.K. 22-25 March 2010.

OECD (2002), *Development Co-operation – 2001 Report*, Paris: Organisation for Economic Co-operation and Development.

OECD (2009), 2009 *DAC Report on Aid Predictability - Survey on Donors' Forward Spending Plans 2009-2011*, Paris: Organisation for Economic Co-operation and Development.

OECD (2011), *The Paris Declaration on Aid Effectiveness and the Accra Agenda for Action*, Paris: Organisation for Economic Co-operation and Development.

Plümper, Thomas and Eric Neumayer (2010), Model Specification in the Analysis of Spatial Dependence, *European Journal of Political Research*, 49(3), pp. 418-442.

Reinhardt, Gina Yannitell (2006), Shortcuts and Signals: An Analysis of the Micro-level Determinants of Aid Allocation, with Case Study Evidence from Brazil, *Review of Development Economics*, 10(2), pp. 297-312.

Round, Raymond and Matthew Odedokun (2004), Aid efforts and its determinants, *International Review of Economics and Finance*, 13, pp. 293-309.

Stokke, O. (ed.) (1989), *Western Middle Powers and Global Poverty – The Determinants of Aid Policies of Canada, Denmark, the Netherlands, Norway, and Sweden*, Uppsala: The Scandinavian Institute of African Studies.

Svensson, Jakob (1999), Aid, Growth and Democracy, *Economics and Policy*, 11, pp. 275-297.

Svensson, Jakob (2008), Absorption Capacity and Disbursement Constraints, in: Easterly, William (ed), *Reinventing Foreign Aid*, Cambridge, MA: The MIT Press, pp. 311-332.

Tarp, Finn, Christian F. Bach, Henrik Hansen, and Søren Baunsgaard (1998), *Danish Aid Policy: Theory and empirical evidence*, Discussion Paper, 98/06, Copenhagen: Institute of Economics, University of Copenhagen.

Torsvik, Gaute (2005), *Foreign economic aid: should donors cooperate?,* Journal of Development Economics, 77(2), pp. 503–515.

USAID (2010), *U.S. Overseas Loans and Grants Online (Greenbook),* available at: http://gbk.eads.usaidallnet.gov/, last accessed: 10.03.2010.

Vázquez, Sergio Tezanos (2008), *The Spanish pattern of aid giving*, Working Paper 04/08, Instituto Complutense de Estudios Internacionales.

Ward, Michael D. and Kristian Gleditsch (2008), *Spatial Regression Models*, London: SAGE.

Woods, Ngaire (2011), Rethinking aid coordination, in: Kharas, Homi, Woojin Jung, and Koji Makino (eds.), *Catalyzing Development*, Washington, DC: Brookings Institution Press.

World Bank (2003), *World Development Report 2004: Making Services Work for Poor People*, Washington, D.C.: World Bank.

World Bank (2010), *The World Development Indicators 2009*, Washington, DC: The World Bank.

Appendix

Table 27: Donor classification

Weighting matrix:	*Same subgroup*
France (1974-2008) Germany (1974-2008) Italy (1974-1977, 1980-2008) Japan (1974-2008) United Kingdom (1974-2008) United States (1974-2008)	Big Western donors
Canada (1974-2008) Denmark (1974-2008) Netherlands (1974-2008) Norway (1974-2008) Sweden (1974-2008)	Like-minded countries
Australia (1974-2008) Austria (1974, 1976-2008) Belgium (1974-2008) Finland (1974-2008) Greece (2002-2008) Ireland (2000-2008) Republic of Korea (1991-2008) Luxembourg (2001-2008) New Zealand (2002-2008) Portugal (1983-1985, 1987-2008) Spain (1988-2008) Switzerland (1974-2008)	Other bilateral donors
EC (European Community, 1974-2008) ASDB (Asian Development Bank, 1974-2008) United Nations (1978-1986, 1988-2008)	Multilateral donors

Notes: Years in brackets indicate data availability.

Table 28: List of recipients

Albania, Algeria, Angola, Argentina, Armenia, Azerbaijan, Bahrain, Bangladesh, Barbados, Belarus, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Cape Verde, Central African Republic, Chad, Chile, China, Colombia, Comoros, Democratic Republic of Congo, Republic of Congo, Costa Rica, Cote d'Ivoire, Croatia, Czech Republic, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Estonia, Ethiopia, Fiji, Gabon, The Gambia, Georgia, Ghana, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hungary, India, Indonesia, Iran, Iraq, Israel, Jamaica, Jordan, Kazakhstan, Kenya, Kiribati, Republic of Korea, Kuwait, Kyrgyz Republic, Lao, Latvia, Lebanon, Lesotho, Liberia, Libya, Lithuania, Macedonia, Madagascar, Malawi, Malaysia, Maldives, Mali, Mauritania, Mauritius, Mexico, Moldova, Mongolia, Morocco, Mozambique, Namibia, Nepal, Nicaragua, Niger, Nigeria, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Romania, Russian Federation, Rwanda, Saudi Arabia, Senegal, Seychelles, Sierra Leone, Singapore, Slovak Republic, Slovenia, South Africa, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Sudan, Suriname, Swaziland, Syrian Arab Republic, Tajikistan, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Uganda, Ukraine, United Arab Emirates, Uruguay, Uzbekistan, Venezuela, Vietnam, Yemen, Zambia, Zimbabwe

4. Spatial Dependence in Asylum Migration

4.1. Introduction

Even though the number of asylum applications in industrialised countries has declined steadily since the early 1990s and reached with slightly over 300,000 in 2009 only roughly one third of their former peak values, there is a strong political rhetoric against asylum seekers. Polls in industrialised nations show that there are strong resentments against immigrants, and albeit voters are better disposed towards genuine refugees, a strong majority supports stronger measures to exclude illegal migrants (ISSP 1995, 2003). This fuels an ever more restrictive asylum policy in many popular destination countries, which is at least rhetorically aimed at bogus refugees, to reduce the number of applicants and to even the highly unequal distribution of the asylum burden across host nations. Asylum migration is widely regarded as a zero sum phenomenon with a fixed number of migrants – if one country deters refugees by curtailing welfare benefits, a tighter visa regime or lower acceptance rates, higher application numbers in other host nations are the consequence.[106] This negative externality has led to political tensions in the past, e.g. the stressed relations between Denmark and Sweden after the introduction of highly restrictive asylum measures by the new conservative government in Denmark as well as the controversy about the Sangatte refugee camp which strained the French-British relations (Thielemann 2006a). More recent examples encompass the controversies about the responsibility for refugees within the European Union after the mass landings in Spain and Italy as an outcome of the political turmoil in Northern Africa and the Middle East or the political huff after the re-introduction of border controls by Denmark.

The standard determinants of asylum migration are well understood after several country case studies and empirical analyses in large scale monadic and dyadic datasets. Many factors such as historical, geographic, economic or partly reputational factors lie beyond the scope of policy makers and leaves asylum policy as one of the main lever.[107] Against the background that many political measures in host countries are taken to detering asylum seekers rather than

[106] Even though asylum seekers constitute only for a small subset of all migrants, both terms are used interchangeably. If not stated otherwise, “migrant” therefore refers to “asylum seeker”. By the same token, the terms “country of origin” and “source country” as well as “host country”, “target country” and “destination country” are synonymous in this text.
[107] However, national asylum policy is constrained by international agreements such as the Geneva Convention on the Status of Refugees.

reducing the causes of asylum migration at the root,[108] policy makers are apparently more interested in lowering the number of applications filed with domestic authorities rather than reducing the number of asylum seekers globally. These deterrence measures together with the same global level of asylum seekers should lead to negative spatial dependence between host countries which are regarded as substitutes from a migrant's point of view. However, theoretical arguments not only suggest spatial dependence between the targets of asylum migration, but also between different source countries. One the one hand, positive spatial dependence between geographically and culturally close source countries could be explained by cross-country network effects which lower the risk and costs of migration. On the other hand, also the organisation of human trafficking can cause positive spatial dependence if asylum seekers from the same region use common people smuggling networks and if minor source countries gain access to these transnational migration routes.

Apart from some anecdotal evidence in qualitative research, this potential spatial dependence in the number of asylum seekers has been fully neglected in the existing literature so far. This work tries to fill this gap and basically seeks to answer two questions: (1) Is the number of asylum seekers from one source country to one destination country influenced by the number of asylum seekers from other source countries to the same destination (specific source contagion)? And (2), does the number of asylum seekers in one target country depend on the number of asylum seekers in other target countries (target contagion)? Hence, the contribution of this paper is twofold: First, it is the first to model the potential spatial dependence both between source countries as well as between target nations. This allows not only checking the robustness of standard determinants of asylum migration which have been established in previous studies to allowing spatial dependence. If such spatial dependence is found, it also has several important policy implications. Second, it is a dyadic study which is based on a larger sample of both source and target countries in the cross-section and that covers a longer time period than any previous work. While the broad country basis minimises the risk of a potential sample selection bias, the long time series covers both the strong increase in the numbers of asylum applications as well as the steady decline since 2001.

---

[108] This is probably at least partly due to the public good characteristics of refugee protection and the opportunity to free-ride (see Section 4).

The remainder of this chapter is structured as follows: Section 2 provides definitions as well as descriptive analyses of trends and distribution of asylum applications. In Section 3, the costs and benefits of asylum migration both for the origin as well as for the host countries are briefly discussed, followed by a brief overview over the main asylum policy levers. Furthermore, the theoretical arguments causing target and specific source contagion are presented. Section 4 reviews the relevant literature, while the dataset in general and the definition of the spatial lags in particular are introduced in Section 5. The main results and some robustness checks are presented in Section 6. Finally, the last section concludes.

4.2. Definition and descriptive analysis of asylum seekers

Article 1 of the 1951 Convention Relating to the Status of Refugees defines a refugee as a person who "owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion, is outside the country of his nationality, and is unable to or, owing to such fear, is unwilling to avail himself of the protection of that country.''[109] The signatories, which encompass all developed and numerous developing countries, obligate themselves to consider any credible application, no matter whether the applicant entered the country illegally or not (Article 31). Refugees enjoy the right of non-refoulement, i.e. the right not to be deported to their country of origin where she or he might be at risk of prosecution (Article 33). The refugee definition of the convention does not directly cover other threats to one's life and wellbeing, such as famine, natural catastrophes, war, and general political violence (Neumayer 2005a). Industrialised countries have always been reluctant to expand the formal definition, mainly because this would entail the duty to provide asylum to migrants affected by these causes (Roberts 1998). An asylum seeker is a person who has filed an application for asylum, but who has not yet received a final decision on his or her application (UNHCR 2011). Asylum seekers can be distinguished from internally displaced persons, who involuntarily have to leave their homes and places of habitual residence due to wars, violence or human- or natural-made disasters without crossing an international

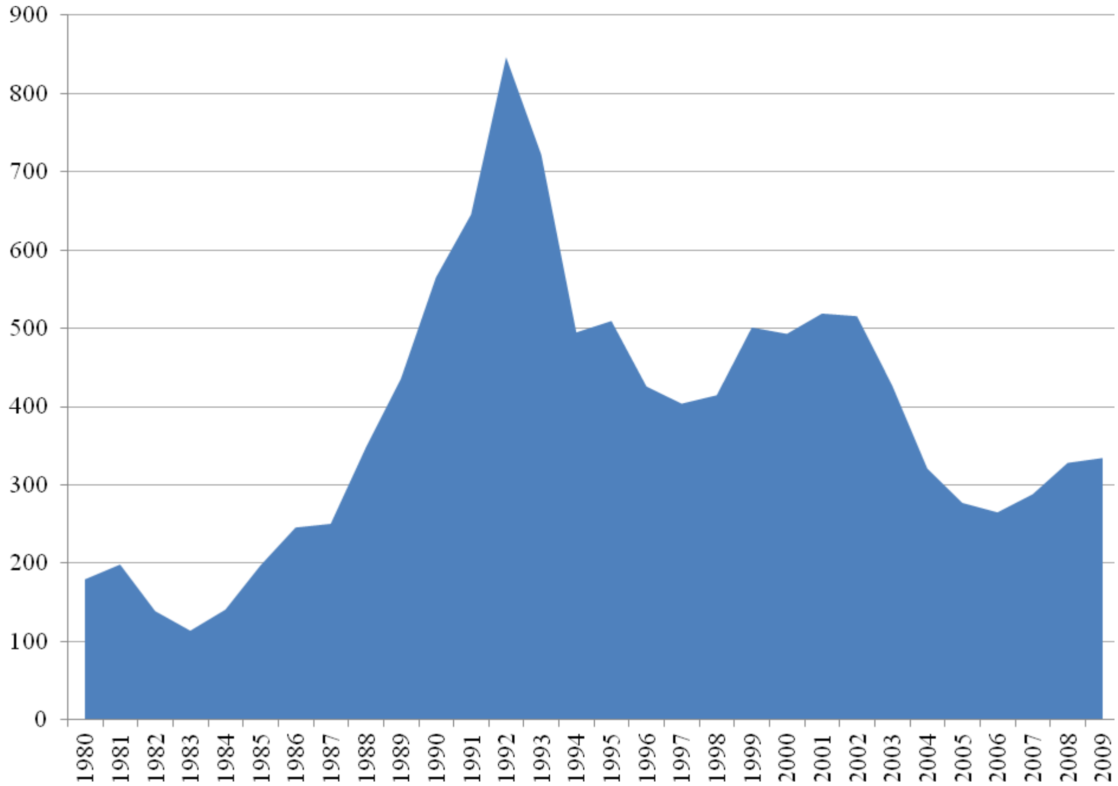
[109] The 1951 Convention was limited to the protection of European refugees after World War II and covered only events occurring before 1951, but geographical and time limits were removed in a 1967 Protocol.

border (UNHCR 2011). As discussed in Section 3 below, there is a plethora of reasons why a person leaves his or her home country to seek their fortune abroad. Not for all forced migrants is one of the treats stated in Article 1 the main reason to leave. Instead, a certain proportion of asylum seekers try to enter the destination country under the guise of personal prosecution, even if the main objective is the sole hope for better living conditions. These bogus refugees are in fact voluntary migrants and the respective administrations in the host countries are anxious to refuse granting asylum to any migrant not in genuine need of protection.

Figure 24 displays the annual number of asylum seekers in industrialised OECD countries since 1980. [110] All numbers refer to first instance applications filed in a given year, rather than to the stock of pending cases. In these countries, the number of applications increased steadily since 1983 and reached the all time high in 1992 with more than 800,000 applications. This surge is mostly due to a large number of refugees as a result of the war in former Yugoslavia, with most of them seeking asylum in Germany and Sweden. The numbers in industrialised countries since then have declined steadily and reached the 1988 level of around 300,000 applications in the latest years.

[110] The following countries are taken as industrialised OECD members: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States.

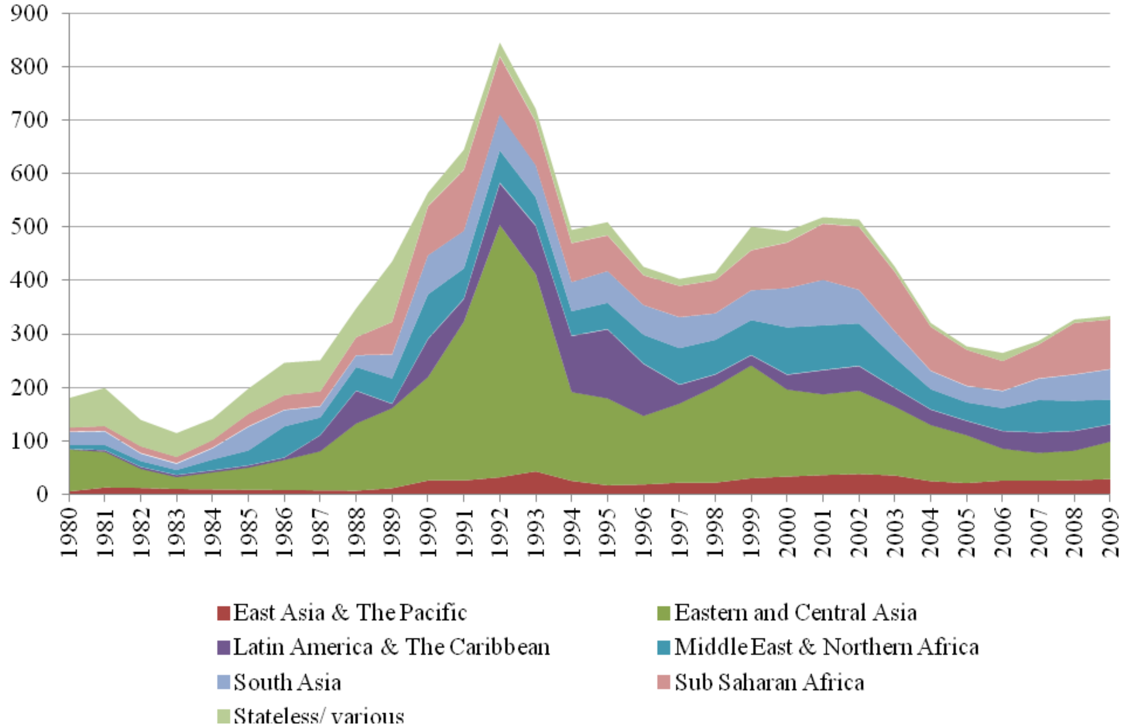
Figure 24: Total number of asylum applications in industrialised OECD countries 1980-2009 (in Thousands)



Notes: Only industrialised countries are taken as OECD members; See Table 35 in the Appendix for a list of target countries in the estimation sample; Data source: UNHCR (2011).

The annual number of applications in OECD countries is broken down by region of origin in Figure 25. Over the whole period, the largest proportion of asylum seekers in industrialised OECD countries originated from countries in Eastern and Central Asia with the countries of former Yugoslavia and Turkey leading the list of major source countries. With the exception of East Asia and the Pacific, the other regions do not differ considerably in terms of their importance as regions of origin over time. The declining number of asylum seekers classified as stateless or with various nationalities is due to better data availability. Some major destinations, such as the United States, do not report the numbers of asylum seekers split up by country of origin for earlier years. The predominance of origin countries in Eastern and Central Asia can at least partly be explained by their geographical proximity to Western Europe, which encompasses 18 out of 23 destination countries in the sample.

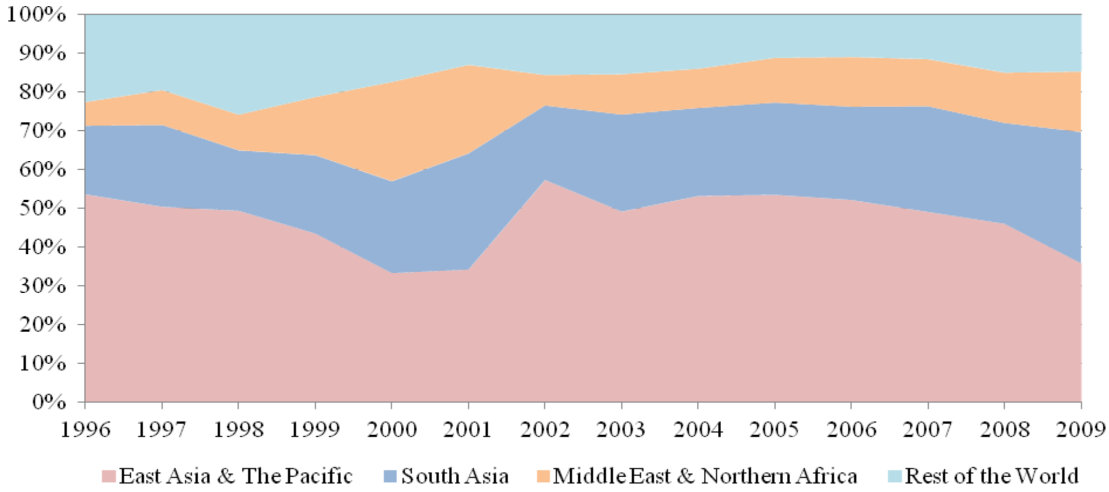
Figure 25: Total number of applications in OECD countries by region of origin (in Thousands)



Notes: Only industrialised countries are taken as OECD members; Data source: UNHCR (2011).

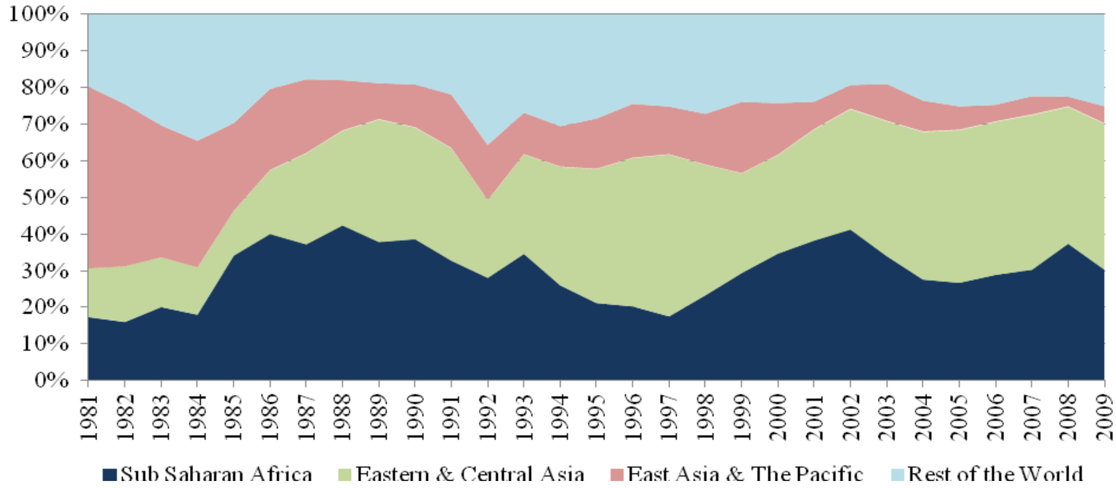
Yet, the relative importance of source regions differs greatly across destination countries as can be seen from Figures 26 to 30, which illustrate the share of asylum seekers from the top three source regions at the total number of asylum applications in different target countries. Nearly half of all applications in Australia are made by migrants that come from countries in the East Asia and Pacific region, followed by around 20 percent of applications filed by people coming from South Asian countries (Figure 26).

Figure 26: Asylum applications in Australia by region of origin



For France, most applications derive from countries in Sub-Saharan Africa, even though Eastern and Central Asia gained relative importance since the end of the cold war (Figure 27). The share of source countries in East Asia and the Pacific, which accounted for 50 percent of all applications in the early 1980s, has decreased considerably to only less than five percent in recent years. The main countries of origin from the Asia Pacific region are Cambodia, Lao People's Democratic Republic and Vietnam, all of which share former colonial ties with France. Not surprisingly, most asylum seekers from Sub-Saharan Africa come from francophone countries such as the Democratic Republic of Congo, Mali and Algeria.

Figure 27: Asylum applications in France by region of origin



For Germany, countries in Eastern and Central Asia have always been the most important sources of asylum seekers (Figure 28), even though Middle Eastern and Northern African source countries and particularly Iraq gained relative importance in recent years. Over the whole period, Serbia followed by Turkey, Romania, and Poland top the list of most important sender countries. In the United Kingdom, the proportion of asylum seekers from Sub-Saharan Africa varies between ten and 40 percent of all applications (Figure 29). In the region, most migrants originate from the former colonies Somalia and Zimbabwe. The second important region in recent years is South Asia with Sri Lanka and Pakistan being the most important sender countries. Both are members of the Commonwealth of Nation. The large share of applicants from the Middle East and Northern Africa in the early 1980s is because of a considerable influx of asylum seekers from Iran, who made up around 60 percent of all applications between 1980 and 1982.

Figure 28: Asylum applications in Germany by region of origin

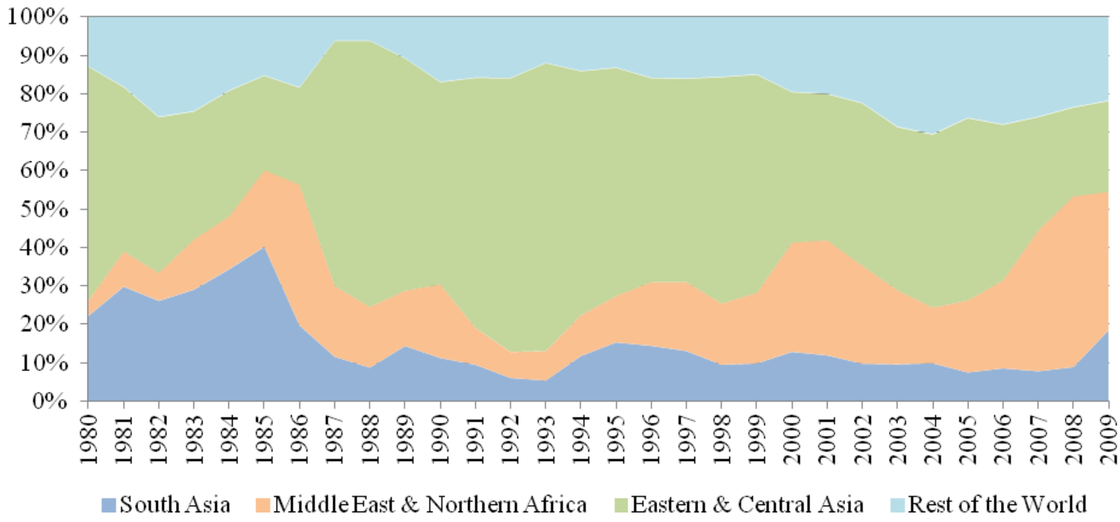
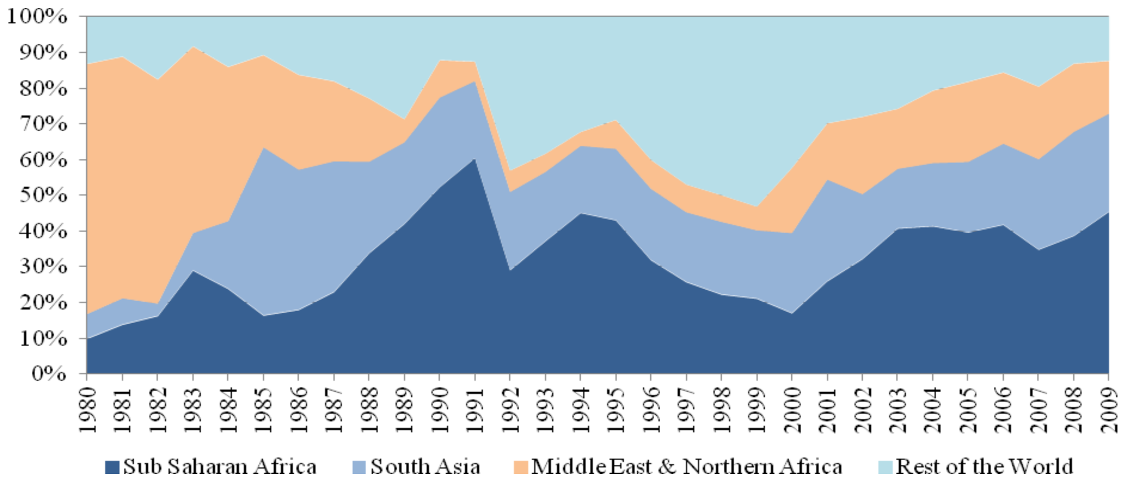
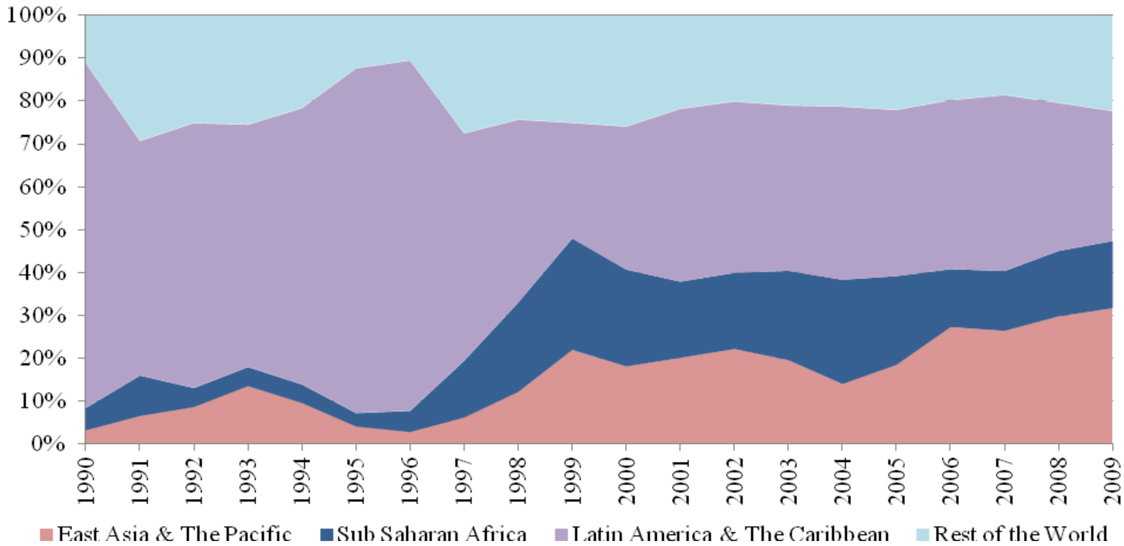


Figure 29: Asylum applications in the United Kingdom by region of origin



For the United States, country of origin specific data of asylum applications is only available since 1990. Over the last two decades, there is a clear dominance of migrants from Latin America and the Caribbean, even if their share decreased in the mid 1990s due to a higher number of applications from East Asia and the Pacific Islands as well as from Sub-Saharan Africa (Figure 30). In the regions, the most important countries of origin are El Salvador and Guatemala, China and the Philippines, and Ethiopia and Somalia.

Figure 30: Asylum applications in the United States by region of origin



These country examples provide two important insights: First and foremost, apparently certain characteristics of the source region – host nation relationship have a great influence on the number of asylum seekers from a region in a destination country. One of these factors is geographical distance which directly influences travel costs. But also other determinants such as a common colonial history, cultural similarity or a common language seem to matter. Second, even though there are some shifts in the relative importance of the source regions, the overall picture is relatively constant over time. On the one hand, this further corroborates the importance of time invariant factors such as distance and common language. On the other hand, this could also be explained by network effects which make asylum migration into a certain country more likely if there is large stock of asylum seekers or migrants from the source region. As discussed in Section 3, the presence of fellow countrymen in the host country reduces the risk of migrating for an individual. The positive impact of such migrant-networks is well documented in the literature (Section 4), but so far it is assumed that these network effects are specific for one country of origin. The constant relative importance of source *regions*, however, points into the direction that these effects could also operate across borders, i.e. that not only a large number of migrants from a specific country of origin facilitates asylum seeking in the target country, but also that the presence of asylum seekers from neighbouring source countries increases the attractiveness of a host country.[111]

But destination countries differ not only with respect to the main source region of asylum migrants, but also considerably with respect to the total number of applications they receive. Any reasonable comparison across destination countries has to take the large differences in size of various target countries into account. Ceteris paribus, a larger country will attract more asylum seekers than a smaller country. In Table 29, the population of the destination country is used to control for the size of the receiving country and the number of applications is expressed as applications per 1000 inhabitants of the host country. Population is by no means the only normalisation factor to make the number of applications comparable across countries, but others such as GDP yield a very similar ranking order. The table reveals considerable differences between countries as well as over time. The last column shows the average number of asylum applications per 1000 people over the whole period. On the top of the list are Switzerland, Sweden, Luxembourg and Austria, which are all in Western Europe and thereby geographically close to the major source regions of Eastern and Central Asia. Even though they differ considerably in population and territory size, they all share a high level of GDP per capita. Generally, lower levels of applications can be observed for island states as these are less accessible. At the bottom of the list are also countries with lower levels of GDP per capita, such as the Czech Republic, Poland or Hungary, but also Italy, Spain and Portugal appear to be relatively unattractive for asylum seekers.

[111] In Section 4, it will be argued that the existence of human trafficking networks might also explain the presence of such spatial dependence in the number of asylum seekers from geographically close source countries.

Table 29: Annual averages of applications per 1000 people in different destination countries

Destination	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2009	average 1980-2009
Australia	n/a	0.08	0.56	0.48	0.40	0.20	0.34
Austria	1.67	1.70	1.96	1.34	3.58	1.84	2.01
Belgium	0.29	0.65	1.73	1.84	2.30	1.29	1.35
Canada	0.23	1.09	1.05	0.85	1.12	0.85	0.87
Czech Republic	n/a	n/a	0.16	0.33	1.02	0.24	0.44
Denmark	0.22	1.65	2.96	1.37	1.39	0.45	1.34
Finland	0.00	0.01	0.45	0.27	0.59	0.65	0.33
France	0.39	0.64	0.65	0.39	0.86	0.61	0.59
Germany	0.64	1.17	3.32	1.62	0.79	0.29	1.30
Greece	0.13	0.48	0.25	0.22	0.49	1.47	0.51
Hungary	n/a	0.00	0.14	0.41	0.55	0.30	0.28
Iceland	n/a	0.04	0.03	0.04	0.24	0.18	0.11
Ireland	n/a	n/a	0.04	0.96	2.33	0.89	1.05
Italy	0.06	0.09	0.14	0.17	0.22	0.27	0.16
Japan	< 0.00	< 0.00	< 0.00	< 0.00	< 0.00	0.01	< 0.00
Luxembourg	n/a	n/a	n/a	2.56	2.44	1.12	2.04
Netherlands	0.12	0.63	1.99	2.23	1.47	0.76	1.20
New Zealand	n/a	0.05	0.19	0.37	0.28	0.07	0.19
Norway	0.04	1.11	1.40	1.08	2.95	2.06	1.44
Poland	n/a	n/a	0.03	0.07	0.15	0.21	0.11
Portugal	0.09	0.02	0.08	0.03	0.02	0.01	0.04
Spain	0.04	0.08	0.27	0.15	0.17	0.12	0.14
Sweden	1.26	2.31	4.55	1.10	2.85	2.77	2.47
Switzerland	0.93	2.14	3.99	4.12	2.73	1.68	2.60
United Kingdom	0.06	0.10	0.52	0.76	1.14	0.48	0.51
United States	0.15	0.18	0.40	0.28	0.16	0.12	0.22
Total number of applications (m)	0.77	1.48	3.29	2.31	2.38	1.56	

Notes: n/a indicates that data is not available; Total number of applications in the last row shows the aggregate number of all applications in the 26 destination countries over each five year period.

Over time, the number of applications generally varies in most countries in line with the total number of applications which are shown in the last row. However, there are also some notable deviations: For instance, while the total number peaks in 1990-1994, Finland saw a constant increase in applications over the whole period, whereas the number decreased in Sweden and Denmark after the high in 1990-1994. Austria faced a considerably higher number of applications in 2000-2004 than in any other period. Also other countries such as Greece and Norway received more applications during the last 10 years than in the previous two decades despite a strong decline in the overall numbers of applications. One possible driver of such variations over time could be war or political unrest in the major source country. Another explanation could be that a stricter asylum policy and lower recognition rates in one destination country deters asylum seekers and motivates them to lodge their application in a more lenient country.

4.3. Costs and benefits of asylum migration for the migrant and the host country, trends in asylum policy and sources of spatial dependence

This section discusses the factors which influence a migrant's decision whether to flee or not and in which country to file the application if the migrant decides to leave the home country. This is done by addressing the costs and benefits of asylum migration from a migrant's point of view. One important factor is the probability that the application in a host country will be successful and asylum be granted. This likelihood depends on the asylum policy in the target country which is in turn influenced by the perceived costs and benefits of asylum seekers from the host country's point of view. Asylum migration can be explained by the fact that the benefits of migration outweigh the costs for some people. An ever more restrictive asylum policy, which can be observed in many industrialised countries and which is outlined in the third subsection indicates, however, that voters and policymakers in many host countries perceive that the costs of a larger inflow of asylum seekers exceed the benefits of hosting them. Based on these considerations, the identification and discussion of the sources of different kinds of spatial dependence is provided in the fourth subsection. The theoretical part forms the basis for the model specification in the econometric analysis.

Costs and benefits for the migrant

Following previous studies (e.g., Moore and Shellman 2003, Neumayer 2004, 2005a), the decision of an asylum seeker to migrate and to lodge his application in a given destination country is modelled as the outcome of a utility maximising behaviour.[112] An individual weighs the costs and benefits of staying in his or her home country against the costs and benefits of leaving and filing an application for asylum in the destination country. This decision is subject to given constraints such as limited financial resources for travelling. The net benefit of staying (leaving) is simply the difference between the costs of staying (leaving) and its benefits. A rational person decides to migrate if the net benefit of leaving exceeds the net benefit of staying, i.e. if migration maximises his or her utility.[113] By the same token, target countries differ with respect to migration costs and potential benefits and therefore the decision where to seek asylum can be directly incorporated into the model of utility maximising behaviour. One has to be bear in mind that most asylum seekers make their decision under great pressure and might not be able to fully balance the costs and benefits of their migration decisions (Neumayer 2005a). However, finally such a decision is made if a person is not literally forced out of a country against his or her will, which is probably true only for a minority of all cases. Even though voluntary and forced migration is based on different main motivations, they share some important similarities. Therefore, some of the costs and benefits addressed below are derived from the literature on voluntary migration.

Sketching the benefits of migration first, economic considerations play a crucial role in explaining migration in general. As argued in neoclassical economic theory, a difference in the rate of return to human capital and hence in the wage rate between the country of origin and the destination country fosters migration. By leaving his country and seeking his fortune abroad, the migrant is able to improve his economic situation. A person is more likely to migrate the poorer the economic prospects in his home country are and the higher the chances to find an adequately paid job in the host country. Low levels of development, little economic growth and

[112] Even though the decision of an individual whether to migrate or not is not subject of this analysis, factors which increase or decrease the probability of an individual to leave his or her home country and ask for asylum in a host nation directly affect the total number of asylum applications from a source in a target country.
[113] As noted by Neumayer (2005), benefits of leaving represent the opportunity costs of staying and vice versa which makes the distinction between costs and benefits obsolete. For illustrative purposes, both the costs and benefits of staying and migrating are outlined separately. The concept is also fully compatible with the push-pull-factor theory employed by some authors (e.g., Thielemann 2006), where negative influences in the home country (push factors) and positive aspects of the destination country (pull factors) foster migration. Since both costs and benefits might be subject to a great deal of uncertainty or materialise only in the future, all values should be taken as expected and discounted present values.

high rates of unemployment in the source country and high levels of GDP per capita and employment as well as a dynamic economy in the potential target country provide the opportunity to improve the livelihood of a migrant.[114] Also, younger people in working age and better educated individuals are more likely to migrate since they are probably more capable of improving their living standard (Borjas 1994) and because for them the net present value of migration is higher (Hatton and Williamson 2005). The chance to escape threats to personal integrity is the second main benefit of migration. Threats stemming from political oppression in the home country are particularly relevant for asylum migration that can find its expression in restrictions on the freedom to assemble, to associate, to compete for a political office or to voice dissident political opinions (Neumayer 2005a). While these discriminatory activities impact all individuals, there are some restrictions which are directly targeted towards members of a specific ethnic or religious group or towards political dissidents. Also armed conflicts, either civil or interstate wars, pose a risk to the life and wellbeing of individuals. Finally, natural catastrophes such as earthquakes, floods, and famines as a consequence of droughts threaten a person's safety. Migration to a safe, politically stable, democratic or less exposed country might be seen as the only way to evade these threats.

Turning to the costs of migration, direct costs of transportation generally increase with the geographical distance between the source and the destination country, but they are also influenced by the availability of air or boat connections or the quality of road and railway infrastructure as well as the route taken, the transport mode and the services supplied by agents.[115] The costs of travel tickets are not the only cost factor for forced migrants as these often have to rely on people smugglers, because they can only file an asylum application from within the host country and destination countries are keen to impede access for potential asylum seekers by imposing visa restrictions (Neumayer 2006). Reliable information on travel costs is generally scarce and estimates scatter widely. Petros (2005) provides estimates for the

[114] Vogler and Rotte (2000) argue that the relative rather than the absolute income situation matters, i.e. a poor among poor has less incentives to migrate if he will end up as a poor among rich, even if the absolute income level would be higher.

[115] Agents basically provide three types of services to asylum seekers (Robinson and Segrott 2000): (1) Provision of travel documents, (2) facilitation of the journey, and (3) channeling the asylum seekers towards particular destination countries. The last point also involves advice about which destination country to chose.

mean costs for a journey from Asia to Europe of around 9,400 USD, from Africa to Europe of 6,500 USD and from the Americas to Europe of 4,500 USD. A report by the UNODC (2011) estimates the costs for Asians taking the West Africa route to Europe at around 12,000 USD. The existence of people smuggling networks often enables asylum migration in the first place, or the costs of smuggling go down as the scope and efficiency of the network increases (Hatton 2004). At least as important – though much harder to quantify – are however the social costs of migration, since one has to leave the familiar environment and relatives behind and adapt to an unknown surrounding. This might not only entail a new language, but also a different culture. In addition, immigrants might not be welcomed by the existing population. The size of these social costs depends on various factors. For example, cultural similarity and a common language between the country of origin and the destination country facilitate transition and lower thereby the costs of migrating. One factor that has been stressed in many scholarly papers (e.g., Rotte et al.1997, Neumayer 2004, 2005a, Thielemann 2004, 2006, Hatton 2009) is the existence of "migrant networks" (Massey 1990). As network theory suggests, a higher number of past migrants with a similar cultural background reduces information, assimilation and transaction costs for potential subsequent migrants and makes migration more likely (Rotte et al. 1997). These personal networks might help to gain a foothold in the host country by assisting the search for accommodation and employment. Individuals who migrated earlier could also be a role-model for potential migrants and their experience with selecting traffickers and obtaining visas might help to reduce the uncertainty necessarily involved in illegal migration (Neumayer 2005a). However, not only personal networks matter; other contacts between the source and the destination countries might also foster migration. Such contacts could be established through trade, tourism or bilateral development assistance (Bilsborrow and Zlotnik 1994). Social costs could additionally occur if asylum seekers are not welcomed in the host country due to xenophobia. This is particularly relevant if these resentments against foreigners express themselves in violent acts or public demonstrations (Neumayer 2004). Finally, also asylum policy and welfare provisions in the host countries influence the costs of migration (Robinson and Segrott 2002). While generous social benefits and work permits for asylum seekers lower the costs, restricted access to employment and to benefit entitlements and dispersal to reception centres increase both the costs and the risk of seeking asylum.[116] Low recognition rates for asylum seekers in the past indicate a tough stance in the host country and are associated with a higher risk of rejection which in turn increases the costs of migration.[117]

Costs and benefits for the host country

Hosting asylum seekers provides some benefits to the host country. First, countries have used asylum policy as a way to promote their own interest. During the cold war, particularly Western European countries welcomed refugees from communist regimes, and granting asylum to them served to "embarrass and discredit adversary nations" (Teitelbaum 1984: 430). As long as the numbers of people seeking asylum was small, the institution of asylum was little disputed and enjoyed the public's sympathy (Neumayer 2005a).[118] Second, humanitarian motives provide an incentive to help others who are in need of protection from threats to their personal integrity. Public opinion is generally favourable to accepting at least genuine refugees who have a well founded fear of prosecution. Granting asylum is therefore satisfying altruistic motives and provides a benefit to individuals of the host country (Hatton 2004). Finally, since asylum seekers are often in working age and tend to have larger families, they form an additional reservoir of workforce (Aldridge and Waddington 2001) and might help to mitigate the problem of an over-aging population that emerges in many industrialised countries.[119] However, since the number of asylum seekers is typically low in relation to the population of the host country (Table 29), these effects are presumably small.

[116] There is a feedback-loop between xenophobia and tight asylum policy if right-wing populist parties which pledge an anti-immigration and anti-asylum policy in their campaigns gain a higher voting share (Neumayer 2005).

[117] Regarding the relative importance of these factors, qualitative research suggests that the choice of a destination country is mainly determined by the presence of friends and relative, whereas policies are of secondary importance (Havinga and Böcker 1999, Robinson and Segrott 2002).

[118] As noted by Hansen and King (2000), the liberal stance changed when the numbers of arriving people soared, refugees increasingly originated from third world countries with a different cultural background than the host countries and when they entered host nations with the help of people smugglers and false documentation.

[119] For a small sample of refugees recognized in the UK in the 1990s, Carrey-Wood (1995) find that one-third held degrees or had postgraduate or professions qualifications and the majority of those had been in professional, managerial or business occupations.

Turning to the costs for the receiving countries, first and foremost, there are the fiscal costs of hosting asylum seekers. A study by Thielemann et al. (2010) commissioned by the European Parliament contrasts the asylum related fiscal costs borne by the 25 EU member states in 2007. Their analysis covers a wide range of different costs factors, such as expenses related to the administration and examination of asylum claims, temporary housing, legal assistance to applicants and the removal of bogus applicants. While the EU-wide overall direct spending accounted for 4.16bn EUR in 2007, individual member states generally spent not more than one fourteenth of their international aid target of 0.7 percent of gross national income on asylum related activities (Thielemann et al. 2010: 17). However, the costs differ greatly across member states and depend not only on the total number of applicants, but also on the national characteristics of the asylum system, such as processing times and the level of benefits provided. Further costs, which are not necessarily asylum seeker specific, are caused by efforts to integrate and help assimilate accepted asylum seekers and other immigrants. Yet, asylum migrants can remain a burden for the social systems even after they have been accepted and granted work permits since their education and skills might not allow them full job market participation. This might be because the qualifications are not recognised, documentation is inadequate or a lacking command of the host country language (Hatton 2004) and has been documented in several country-studies. For example, Krahn et al. (2000) show for a sample of refugees accepted in Canada that the probability of un- or underemployment is higher for this group than for nationals. Also, even if many of these immigrants are well educated and worked in managerial positions prior to their flight, it is difficult to re-enter the job market in an adequate position. For the Netherlands, Hartog and Zorlu (2009) find that refugees have lower levels of education than other immigrants and even after five years, only one third was in employment. In Denmark, a gap in the employment rate of prime age male refugees and other immigrants is observed and even the difference diminishes over time, a wage penalty remains (Husted et al. 2000). Also in the UK, employment rates of refugees are considerably lower than of other members of ethnic minorities with similar education (Bloch 2002). However, not only monetary costs occur as a consequence of a large influx of asylum seeking migrants since xenophobia might foster anti-immigrant public opinion and support to populist right-wing political parties as citizens may worry that immigrants threaten their national, ethnic and cultural identities (Leblang 2010).[120] These resentments could result in tensions and threaten internal peace.[121]

Policy measures to limit number of asylum applications

Since the costs of hosting asylum seekers generally exceed its benefits at least in the point of view of voters and policymakers in destination countries, over the last three decades several adoptions to asylum policy were made, mostly with the aim of reducing the number of asylum applications.[122] Besides unilateral policy changes, there are some multilateral approaches within the European Union to harmonise asylum policy regulation across Europe. A comprehensive review of these policies, however, is beyond the scope of this work. Instead, only the key developments and policy levers will be sketched. A more detailed discussion of asylum policies in industrialised countries is among many others provided by Hailbronner (2000), Schuster (2000), Zetter et al. (2003), Thielemann (2004), Gibney and Hansen (2005), and Hatton (2004, 2009).

Hatton (2004, 2009) categorises unilateral asylum policy levers into three main types:[123] (1) conditions relating to the access to the territory, (2) conditions relating to the processing of applications and the determination of status, and (3) conditions relating to the welfare of asylum seekers. Since the mid 1980s and particularly since the mass influx of asylum seekers into Western Europe in the early 1990s, asylum policy in most industrialised countries became more and more restrictive with the main aim of limiting the number of applications.[124] The first

---

[120] A cross-sectional study by Facchini and Mayda (2008) which covers up to 34 advanced economies shows that countries in which voters are critical towards immigration implement more restrictive immigration policies and hence experience less inward migration.

[121] In the US conflicts are predominantly between newcomers and other minorities, whereas they occur between newcomers and the majority population in Europe (Martin et al. 2005).

[122] As argued by Thielemann (2006a), the effectiveness of deterrence policies depends on the assumption that asylum seekers are well informed about the asylum regimes and the differences across destination countries as well as that asylum policy is a major determinant of their choice where to file an asylum application (in particular that refugees are attracted by countries with higher acceptance rates and more welfare provisions for asylum seekers).

[123] Gibney and Hansen (2005) further distinguish policy measures that deter arrivals and measures to limit the stay of asylum seekers.

[124] The question, whether asylum policies are subject to spatial contagion is an interesting one for itself. If the main objective of asylum policy is to deter asylum seekers to other host countries, governments have an incentive

group encompasses for example visa requirements, the possibility to apply for asylum from abroad, the strictness of border controls and the designation of airports as international zones. Also, penalties for agents who smuggle illegal migrants into the country and for carriers by land, sea and air of undocumented arrivals fall into this category. While tighter visa requirements affect not only refugees but all migrants, they are frequently used as a method to impede the inward migration of asylum seekers as they are often introduced for nationals of countries deemed to produce large numbers of asylum seekers (Gibney and Hansen 2005). For instance, the Austrian authorities raised the visa requirements for Poles in 1981, the French for Algerians in 1986, and Canada for Hungarians in 2001. Particularly after the implementation of the Schengen Agreement, which eliminated border controls between its signatory states in the European Union, visa restrictions were progressively tightened up (Hatton 2009). Carrier sanctions also have been introduced by a large number of potential host countries, such as Greece in 1992, Austria in 1997, and Ireland in 2003. By the same token, compliance with policies against human traffickers increased in many regions of the world over time (Cho et al. 2011). In order to avoid the obligations to provide asylum seekers with the protections available to persons who are officially on state territory, some states (e.g., Switzerland, France, Germany, and Spain) declared parts of their airports international zones (Gibney and Hansen 2005). The second group of policy measures includes the definition of a refugee, the speed of processing the application and the possibility to appeal against a decision. Another example of policy measures which influence the determination of status is the designation of countries as "safe third countries" or "safe countries of origin" (Neumayer 2004). This safe country concept was integral part of the Dublin Convention, which became effective in 1997 and was signed by many Western European countries. However, it already was widely applied before and asylum

to ever tighten policy measures in order to compensate their disadvantage of being a relatively more lenient destination which could lead to a higher number of asylum seekers. This strategic interaction between host governments could then lead to a race-to-the-bottom (Thielemann 2006). Due to its focus on a dyadic framework, this question is beyond the scope of this paper. The empirical confirmation of such a strategic interaction has yet to be tackled; however, two empirical studies analyse such spatial policy contagion in the broad field of international migration: Cho et al. (2011) look at spatial contagion in policies against human trafficking, while Brücker and Schröder (2010) analyse how immigration policy of a country is affected by policy choices of other countries. The former find that policies against human trafficking diffuse across neighbouring countries and main trading partners as well as among countries which share political and cultural similarities. The latter find that skill based immigration policies spatially depend on the immigration policies of neighbouring countries and of countries which might be regarded as a substitute form a migrant's point of view, as defined as geographically close and countries with a common culture or language.

was denied for migrants who originated from or travelled through such a safe country. Finally, the welfare conditions of asylum seekers are determined by the extent to which they are allowed to work in the host country, whether the benefits are granted in cash or kind, whether they are free to choose their place of living or whether they are allocated to detention camps. In addition, the possibility of family reunification after a successful application and the chances of being deported after an unsuccessful application impact the welfare of asylum migrants. Limitations on employment are a frequently used tool to protect the domestic labour market and to discourage economic migrants trying to exploit the asylum system. For instance, France withdrew the permission to work for asylum seekers in 1991, Germany followed suit in 1997 and the United Kingdom in 2002. While some countries, for example Germany and Belgium, provide benefits to asylum seekers in-kind rather than cash payments, others such as the United States, France, and Italy entitle asylum seekers to less welfare benefits than permanent residents or entitlement is subject to stricter conditions (Gibney and Hansen 2005). Also housing asylum applicants in detention centres is a measure to make a host country relatively less attractive. Often, these centres are located far away from major cities in outback areas. In Germany, asylum seekers are allocated to such detention centres and they are only allowed to move freely within the county (Landkreis) of their accommodation. The UK introduced a nation-wide dispersal system for asylum seekers in 2000 to work against the concentration of asylum seekers in certain metropolitan areas (Thielemann 2004).

These unilateral policy measures are sometimes accompanied by bilateral treaties. Of particular importance are readmission agreements in which two countries agree to mutually readmit – without any formality – their nationals who reside in the other contracting state without authorisation. This also covers asylum seekers whose application has been rejected on whatever grounds. Not surprisingly, these treaties are regularly concluded between important source countries and major destination countries. For example, the Netherlands signed repatriation treaties with Sri Lanka, Somalia, Ethiopia and Angola in 1997, Spain implemented such a program with Bulgaria in the same year and Sweden entered treaties with Yugoslavia, Bulgaria, Poland, but also with Denmark and Germany, in 1998. While deportation of rejected asylum seekers is a rare phenomenon due to moral, practical and financial constraints, the existence of a bilateral treaty of this kind should discourage asylum seekers with low chances of recognition to migrate and seek asylum.

Multilateral approaches in asylum policy are concentrated in the European Union and emanate from the perception that asylum seekers are a burden for host countries which needs to be shared equally among members.[125] Three concepts of burden-sharing can be distinguished: First, physical redistribution of asylum migrants. In 1992, Germany proposed a physical reallocation of asylum seekers to EU members according to a formula that gives equal weight to population, territory size and GDP. However, since this initiative found no support, movement of asylum seekers is based on double-voluntarism of both the applicant the receiving country (Thielemann 2004). The second approach is based on financial compensation and fiscal redistribution to member states which experience a disproportionately large number of asylum applications. The European Refugee Fund (ERF) was established in September 2000 and supports projects for reception, integration and repatriation of refugees and displaced persons, but its impact is little due to limited financial endowment (Vink and Meijerink 2003).[126] The third pillar of burden-sharing is policy harmonisation. Since the 1980s, EU members have worked towards a convergence of national laws on forced migration. The Dublin Convention, signed in 1990 by twelve original signatories, establishes rules for the country which is responsible for dealing with the application of an asylum seeker, which is usually the state of first entry.[127] This regulation tries to avoid asylum-shopping, where an asylum seeker lodges applications in several countries in order to exploit differences in the

[125] Closely related to the concept of burden-sharing is the question whether the relative number of asylum seekers in EU member states has converged over time, i.e. whether the inequality in application numbers across countries decreased over time. The empirical evidence for Western Europe is ambiguous: Vink and Meijerink (2003) find a convergence over time, but this general trend is interrupted by peaks in disparities in certain years with large inflows. Thielemann (2004) finds the same spikes in crisis-driven disparities, but no overall trend; whereas Neumayer (2004) concludes that there is no sign of convergence over the period 1982-1999, but a weak tendency of decreasing inequality during the 1990s. This suggests that the results are somewhat sensitive to the method and sample.

[126] For the period 2008-2013, the fund has financial resources of 614 million Euro compared to an EU-wide direct spending for asylum related activities of more than 4bn Euro in 2007 alone (Thielemann et al. 2010: 17).

[127] This responsibility of the country of first entry means that asylum seekers who travel overland are not longer able to file the application in the destination country of their choice, but that the country through which they entered the European Union is in charge of their case (Thielemann 2004). At the same time, the burden is shifted from destination countries in the core of the EU to countries sharing a border with a non-member country, mainly to countries in Eastern and Central Europe (Noll 2000).

nation states' welfare provisions. At a ministerial meeting in London in 1992, a consensus on three other issues was reached (Hatton 2005): The safe country of origin and safe third country concept discussed above as well as the concept of manifestly unfounded asylum applications, for which an expedited refugee status determination without the right to appeal was agreed. Other important steps were the 1995 Resolution on Minimum Guarantees for Asylum Procedures, the 1999 Amsterdam Treaty which established a common European Asylum System, the 2002 Agreement at the JHA Council in Brussels regarding a common definition for persons eligible for refugee and subsidiary protection, and the directive on common reception procedures in 2003 (Thielemann 2004).

Critics of these efforts to converge and harmonise asylum policy argue that the main result is not shifting the burden between the member states but onto third countries and the home countries of asylum seekers (Roberts 1998). Rather than the alignment of principles, the outcome is that standards have been lowered in a concerted manner (Holzer and Schneider 2002). While these political initiatives at least ostensibly try to equalise the burden between member states, the European Agency for the Management of Operational Cooperation at the External Borders of the Member States of the European Union (FRONTEX), founded in 2005, is a multilateral institution with one of the main objective being to secure the European Union's external borders. This frequently involves protecting the fortress Western Europe against refugees, by tighter border controls, joint patrols and push-backs on sea. Especially the operations of FRONTEX in the Mediterranean Sea have been heavily criticized as blocking persons in need from seeking asylum, driving them onto even more dangerous routes and forcing them to resort to the services of dubious people smugglers (UNHCR 2010).

Summing up, on the one hand the net benefit and the risk an individual faces by asylum migration is influenced by a multitude of factors. The large number of people who leave their home country and lodge an asylum application in a remote country indicates however that there is a perceived positive net benefit for many. On the other hand, an ever tighter asylum regime in Western countries and efforts to curtail the number of asylum applications suggests that these main destination countries weight the costs of hosting asylum seekers larger than the benefits. Both on the source and the target side there are theoretical reasons why one would expect spatial dependence, i.e. why the number of asylum seekers from a given source country is not independent from the number of asylum applications from other countries, just as the number of applications in a target country is not independent of applications in other target countries.

Sources of spatial dependence

Asylum migration is a classical example of a dyadic phenomenon, which involves two countries. More specifically, a country-pair is an example for a directed dyad where the flow originates in the home country of the migrant *i* and is directed to the destination country *j*. Neumayer and Plümper (2010) provide a categorisation of all possible forms of spatial dependence in such dyadic data. Spatial dependence in undirected dyads notwithstanding, five different forms of spatial dependence can be distinguished: Under directed dyad contagion, the number of asylum migrants in $dyad_{ij}$ depends on the level of asylum migration between other dyads *km*. This would simply measure whether there is a global trend in the number of asylum seekers. Aggregate source contagion exists, if the number of asylum seekers in $dyad_{ij}$ depends on the aggregate number of applicants from other sources k ($\neq i$) to all other targets *m*, not just to target *j*. Similarly, the number of asylum seekers in $dyad_{ij}$ depends on the aggregate number of applicants in other targets *m* from other sources *k*, not just the specific source *i*, under aggregate target contagion. The two remaining forms are specific source contagion and specific target contagion. Specific source contagion means that the number of asylum seekers from source *i* to target *j* depends on the weighted number of asylum seekers from other sources *k* to the very same destination country *j*, whereas specific target contagion exists if the number of applicants from source *i* to target *j* depends on the weighted sum of asylum seekers from the same source country *i* to other targets *m*. In the following, the theoretical reasons why one would expect the existence of specific source contagion as well as specific target contagion are discussed. For example, under specific source contagion, the number of asylum seekers from Ghana in the United Kingdom would depend on the number of asylum seekers from Nigeria in the UK. Under specific target contagion, the number of Ghanaian asylum seekers in the UK depends on the number of applications from Ghana in France.

Basically, there are two main theoretical arguments for the existence of positive specific source contagion, namely cross-border network effects and the design of international people smuggling networks. Migration networks, i.e. personal relationships with other migrants with a similar cultural background, ease transition and acclimatisation in the host country. Services provided by a personal network encompass lower search costs for accommodation and work, a sense of belonging and the chance to uphold the cultural habits within the diaspora community. The importance of these networks as a determinant of the size of bilateral asylum flows is well documented in the empirical literature (e.g., Rotte et al. 1997, Vogler and Rotte 2000, Neumayer 2004, Hatton 2004, 2009, Leblang 2010). However, all existing scholars assume that these network effects only exist for migrants from the very same country of origin as the asylum seeker himself, even though Beine et al. (2011: 10) note that "restricting the diaspora to people with the same nationality might be restrictive". This is not necessarily the case if the presence of migrants with a similar cultural background facilitates asylum migration and this similarity is not limited to fellow countrymen as ethnic networks do not necessarily correspond to national borders (Beine et al. 2011). For instance, if there is not only a Ghanaian community but a West African community in the UK which supports individual asylum seekers from the region, more migrants from Togo and Cote d'Ivoire in the UK should reduce the risk and costs of migrating from Ghana and therefore foster asylum seeking by Ghanaians in the UK.[128] One would expect however that these network effects decay relatively quickly with distance and are most pronounced among countries who share a common border. Network effects could also be based on a same language between immigrants, e.g. a francophone community. If it is only a common language with other immigrates which facilitates integration in the host country, the network effect should not decrease over distance between the source countries, i.e. migrants from Haiti in the UK provide the same benefit for asylum seekers from Senegal as do migrants from Guinea.[129]

---

[128] Also, a negative effect is thinkable if there is a rivalry between compatriots of neighbouring countries. In this case, the presence of migrants from one country in a host country would deter asylum seekers from another country. This is reflected in a negative sign of the spatial lag.

[129] Of course, the effect of cultural similarity and common language cannot be disentangled. Hence, the benefit of other migrants from Guinea might be larger than the benefit provided by migrants from Haiti.

The use of services provided by agents is a wide-spread phenomenon among illegal migrants: According to van Moppes (2006), 97 percent of 5,836 irregular migrants of all origins in the Netherlands received assistance from professional people smugglers. The International Centre for Migration Policy Development (ICMPD) estimates that Europe-wide around half of all illegal migrants rely on human traffickers (van Moppes 2006). The range of their services does not only encompass the provision of forged travel documents, such as passports and visas, but also transportation services. During the 1980s, the easiest way for a relatively small group of migrants in Europe was to obtain a visa, arrive by plane, make a false claim for political asylum and go underground if the application was rejected (UNODC 2011). However, this route was closed in the 1990s after the introduction of stricter visa regimes. As a consequence and due to an ever growing number of illegal migrants, overland people smuggle routes developed. There are well-established main transport routes from Central and Northern Africa as well as the Middle East to Europe. These itineraries are increasingly being used by migrants from South Asia, too (UNODC 2011). There are also routes in place to enter the EU from the north trough Russia and the Balkan, from the east trough the new member states and from the south via Albania (International Organization for Migration 2000). While these people smuggling networks often enable illegal migration in the first place, increasing scope and efficiency lower the costs of migration (Hatton 2004). If such a route is established between a major sender country of illegal migrants and a host country due to high demand, it is cheaper for surrounding and transit countries to get connected to the network rather than setting up a bilateral route to the recipient country independently. This is particularly true for overland routes, but also for feeder routes to major airports and harbours from which carriers to industrialised countries leave. Since migration costs are influenced by the existence and capacity of the people trafficking networks, a highly frequented - and thereby more cost-efficient and secure - route might attract migrants from minor source countries. Hence, the size of the asylum flow between close-by or transit source countries and a target country is positively influenced by the number of asylum seekers from a major source country. The assumption of a joint use of trafficking routes by migrants from different countries is corroborated by the presence of various nationalities on the refugee boats that frequently land at the shores of the Canary Islands, Lampedusa or Malta (The International Herald Tribune 2011).

The main source of target contagion stems from the well-known problem of collective action (Olson 1965) where host countries opt for free-riding at the expense of other host nations in the case of providing asylum for refugees. As argued by Suhrke (1998), refugee protection has characteristics of an international public good, where increased security can be regarded as a non-excludable and non-rival benefit, since refugees alleviate the pressure from fuelling and spreading the conflict they are fleeing from (Thielemann 2006). As outlined above, Western policy makers use a multitude of policy levers to render their country less attractive for asylum seekers, such as restrictions on welfare benefits and employment opportunities, increasing the risk of being rejected due to low recognition rates, providing limited opportunities for appealing against a decision and the raising threat of forced removal. As a consequence of this free-riding behaviour, potential host countries' governments have an incentive to engage in a race-to-the-bottom in asylum standards, both concerning welfare provisions and deterrence measures. If tighter asylum policy in one host country deters asylum seekers from filing their application in this country and leads to a higher number of applications in other countries, this is a classical example of a beggar-thy-neighbour policy (Rotte et al. 1997). At the same time, it results in the number of asylum applications in one country depending on the asylum applications in another country or as Böcker and Havinga (1998: 263) conclude: "[...] the introduction of measures to reduce the influx in one country may produce rising numbers in neighbouring countries". To which country asylum flows are deflected depends on which countries are regarded as substitutes from an asylum seeker's point of view. In general, asylum seekers could be incentivised to lodge their asylum application in any more lenient country. While this question has not been analysed empirically, anecdotal evidence suggests that close-by and mostly adjacent countries are frequently seen as an alternative: According to Brochmann (1995) for example, the 1993 introduction of visa requirements for refugees from Bosnia-Herzegovina in Sweden and Denmark led to a marked increase in the number of refugees to Norway. Schilling (1995) notes Denmark and Norway as well as Germany and the Netherlands as two other country pairs with a close substitutive relationship. Böcker and Havinga (1998) point out that the large drop in asylum applications in France in the early 1990s was due to asylum seekers being directed at other countries, e.g., Turkish and Vietnamese asylum flows to Germany, the Angolan flow to the UK and the Peruvian flow at Spain. As a reason for this diversion, a number of policy changes and lower recognition rates in France are adduced, such as the negative asylum decision for many Turkish and Angolan asylum seekers and the introduction of transit visas for Angolans. Holzer et al. (2000) argue that Germany and Switzerland are close substitutes for potential asylum seekers. Rotte et al. (1997: 109) analyse the impact of changes in the French asylum policy on the number of asylum seekers in Germany, as the former is "the major alternative to Germany in Europe for asylum seekers". A successful diversion of asylum flows to other destination countries would lead to negative spatial dependence among target countries. However, also a successful burden-sharing between host nations is in line with a negative coefficient of the spatial lag.[130]

If asylum seekers react to policy changes, the number of applications in a host country will decrease. And if refugees are not fully discouraged to leave their home country but to seek asylum in another industrialised country, this provides an argument for specific target contagion, where the total number of applications from source *i* in country *j* depends on the number of applications from country *i* in target *m*. This is the case if restrictions in asylum policy are targeted against refugees from a specific source country (e.g., visa restrictions or lower recognition rate). Such a targeting of deterrence measures is not uncommon as there are major source countries for certain host countries and natural or human crises can trigger a mass inflow from a specific country (Section 2).

4.4. Literature review

Work on the international flows of people has mostly focussed on migration in general or refugees in particular,[131] but there are still few comprehensive empirical studies on the factors which determine the number of asylum seekers, even though this topic has gained popularity over the last decade. These studies can be broadly classified into monadic and dyadic studies: while the latter explicitly model asylum seekers from a source to a target country, monadic studies use aggregate number of asylum applications in (from) a country without distinguishing individual source (target) countries. However, with the notable exception of Rotte et al. (1997)

[130] See Section 5 for the empirical operationalisation and the concept of spatial lags.

[131] See, for example, Vogler and Rotte (2000) for the determinants of international migration and Schmeidl (1997), Davenport et al (2003), Moore and Shellman (2004, 2006, 2007) on the determinants of the number of refugees and internally displaced persons.

and Hatton (2004) so far no study has explicitly addressed the potential spatial dependence in asylum migration outlined above and even these studies look at the effect of foreign asylum policy and foreign economic variables on domestic asylum applications rather than at spatial dependence in the flows itself.

Monadic analyses can further be divided into cross-country studies, longitudinal country studies and panel studies on the one hand as well as into studies looking at source country determinants and host country characteristics on the other hand. Cross-country studies compare the number of asylum applications in various countries at one point in time and evaluate how differences across these destinations help to explain the distinct attractiveness of a country for asylum migrants. In a simple correlation analysis with averaged data for 17 European countries over 1985 to 2000, Thielemann (2004) finds that GDP per capita in the host country, the share of foreign born population, the size of the ODA budget and a tight asylum policy are positively correlated with the number of asylum applications per 1000 people, while higher unemployment and higher distance to major source countries are associated with lower numbers of asylum seekers. The positive correlation between political deterrence measures and the number of applicants is at odds with expectations and is interpreted as a signal for the ineffectiveness of these policies to deter migrants from seeking asylum in a given destination country. However, the positive correlation could also be due to the potential reverse causality problem as these deterrence measures might also be a consequence of a strong influx of asylum seekers. A major drawback of this analysis is that each factor is analysed individually, which neglects the potential correlation among these determinants.

From an empirical point of view, cross-sectional studies suffer from an important drawback, since it is usually impossible to fully control for all factors that influence the number of asylum seekers in a country. These unaccounted influences are subsumed in the error term. If these factors are correlated with one or more of the control variables, this leads to biased coefficient estimates of these control variables. Furthermore, such studies can only explain the asylum applications in one year, but not the changes over time and their causes. These limitations are overcome by longitudinal country-studies. Here, the inclusion of country-fixed effects allows controlling for all unobserved factors which do not vary over time. For the case of Switzerland, Holzer et al. (2000) find that a tighter asylum policy deterred potential applicants to seek asylum in the Alpine state during 1986 to 1995. The same is true for lower recognition rates.

While these studies overcome the limitations of cross-sectional empirical analyses, they are constrained to one destination (or source) country. This restriction is prone to a severe sample selection bias if the country under scrutiny is not representative for other potential hosts or sources. Panel-studies unite a larger country-sample in the cross-section with a time dimension and exploit the benefits of both approaches. Vink and Meijerink (2003) find a negative correlation between the recognition rate and the numbers of applicants in 10 out of 15 analysed EU member countries over 1982 to 2001. The same negative relationship is established over time, which is interpreted as a political reaction to increased burden. However, the study does not control for other important domestic policy and economic variables which might also influence the number of asylum seekers in a country. If these factors are correlated with the recognition rate, the estimates of the effects are biased.

The first comprehensive monadic study encompassing a large number of source countries is by Neumayer (2005a) who analyses a large number of potential push factors. The dataset covers the aggregate number of asylum seekers from 127 source countries to 15 Western European countries over the period from 1982 to 1999 – a time span marked by a tremendous increase in the number of asylum applications. Estimating both a fixed effects model and a random effects model, he finds that the number of asylum seekers is lower from countries with a higher level of GDP per capita and a faster growing economy. At odds with expectations, the same is true for countries which share a colonial history with any Western European country and for countries with a higher number of tourist arrivals. Both effects, though statistically significant, are negligible in size. The list of factors that increase the number of asylum seekers from a country includes the total population of the source country, the stock of past asylum seekers, economic discrimination and violation of human rights, the share of population aged between 15 and 64, an autocratic regime, the occurrence of violent dissident political activity, as well as both domestic and external wars. Furthermore, a higher number of asylum seekers originates from countries which are closer to Western Europe. In contrast, no significant effect can be established for the numbers of deaths from genocide and politicide and of natural disaster victims, the level of food production, the share of urban population, and cultural similarity as measured by the share of Christians in the source country. Furthermore, both aggregate aid and trade volume seem not to influence the number of asylum seekers.

Focussing on target country characteristics in a large monadic dataset, Thielemann (2006a) seeks to explain the number of asylum applications per capita of the destination country. The sample covers 20 OECD destinations over the period 1985 to 1999. While low levels of unemployment seem to attract asylum seekers, no such effect is found for economic growth in the host country. Network effects, as measured by the stock of foreign population from the top five asylum source countries in a year, prove to be a highly significant determinant of the relative number of asylum applications.[132] A larger ODA budget is also associated with a higher share of asylum seekers. This is interpreted as a sign for a more liberal reputation of the destination country which is appealing for asylum seekers. Generally, a deterring effect of a more restrictive asylum policy is established; however, not all deterrence policies seem to have the same negative impact: whereas not allowing asylum seekers to work while their application is pending and granting protection to a smaller number of asylum seekers discourages migrants to file an asylum application, the safe third country provision, restricted freedom of movement within the host country and providing benefits in kind rather than in cash seem to have no limiting effect. Thielemann (2006a) argues that this is because information about employment opportunities and status determination reaches asylum seekers through human traffickers, personal networks and agents, whereas details about single policy measures might not be available to potential asylum migrants or could be less relevant for the decision to seek asylum.

Hatton (2009) examines asylum migration using both a monadic as well as a dyadic dataset. The monadic study focuses on source country characteristics and seeks to explain the number of asylum seekers per thousand of source country population. Both a country-level random effects and a fixed-effects model are tested. Depending on model specification, he finds that an additional thousand dollars in GDP per capita is associated with a 15 to 25 percent lower share

[132] The inclusion of the stock of asylum seekers from more than one source country is consistent with the concept of specific source contagion. However, Thielemann (2006a) only includes the top five source countries without modeling the relationship between them. Furthermore, the stock of asylum seekers is not host country specific.

of asylum applications. Similarly, fewer applications come from countries with a democratic system and more political rights for its citizens, whereas a higher number of battle deaths increase the demand for asylum. However, controlling for political terror such as arbitrary imprisonment, torture and political murders weakens the coefficients on democracy and political rights and eliminates the effect of a war. The effect of political terror is substantial as well: a one point increase on a scale from 1 to 5 is associated with 50 to 69 percent more asylum seekers.

The question whether official development assistance (ODA) reduces the demand for asylum is in the focus of a study by Leblang (2010), who also exploits both a monadic and a dyadic dataset. The monadic model explains the aggregate number of asylum seekers in 125 source countries over the period 1980 to 2007. In line with previous research, fewer asylum seekers originate from richer countries, whereas curtailed human rights, ethnical wars and natural disasters increase the number of asylum seeking emigrants from a country. Little support is found for the hypothesis that a lower level of democracy acts as a push factor. While the total overall development assistance (ODA) as a share of origin's GDP is associated with lower asylum seekers from a given country, no significant results are obtained if humanitarian, governmental and post-conflict aid are estimated individually. However, as argued by the author, this does not necessarily represent a causal effect as aid is correlated with other factors that might mitigate the causes for asylum seeking or measure the impact of ODA on the direct causes for asylum.

The empirical work by Hatton (2004) is based on a kind-of dyadic setting, in which the number of asylum applications from one of three regions (Africa, Asia and Eastern Europe) divided by the population of the source region is explained. A set of source region variables is added to control for regional push factors. In line with most monadic studies, the author finds that for the period 1981 to 1999, a higher GDP per capita difference, fewer political rights and lower civil liberties foster asylum migration. The same positive effect is found for the stock of source region nationals in the destination country and for the lagged cumulative numbers applications from the source region. By contrast, a more restrictive asylum policy in the destination country lowers the number of asylum applications. A dummy taking the value of one for asylum flows from Africa and Asia to Italy from 1990 on has a negative and highly significant impact. This accounts for the fact that Italy acted as a transit country prior to 1990 for asylum seekers which were transferred onwards to other destinations, as Italy only recognised refugees from other European countries. This reservation was abandoned in 1990 and lead together with a sharp crackdown on asylum seekers to a considerable decrease in the number of applications. A second dummy for asylum applications from Eastern Europe after the end of the cold war in 1989 is positive and accounts for the substantially increased number of asylum seekers in the aftermath of the fall of the Berlin Wall. The average of the asylum policy variable for all other destinations in the sample is positive, yet insignificant. This points into the direction that a more restrictive policy towards asylum seekers in other countries increases the number of applications in a given host country.

The first real dyadic dataset is exploited by Rotte et al. (1997) who analyse source country determinants of asylum outflows in a sample covering 17 sources and Germany as the only destination over the period 1985 to 1994. In a pooled OLS and a random-effects estimation, they find that less political freedom and political terror increase the number of applications in Germany as does a higher living standard and a larger difference between domestic and German GDP per capita. The same is true for the number of nationals already residing in the host country and for German exports into the source country. Asylum reforms which generally result in a less lenient policy discourage asylum seekers to file their application in Germany. To test for a potential substitutability of Germany and France as targets from the point of view of asylum seekers, Rotte et al. include measures for French asylum policy, the relative economic situation between Germany and France and bilateral political relations of France with the source country. No evidence is found that a tighter asylum policy in France increases the number of asylum applications in Germany. Also, a relative higher unemployment in France than in Germany does not encourage asylum seekers to file for asylum in the latter country. However, if France has a strong cultural ties with a given source country, fewer applicants come to Germany from these nations. While this analysis is definitely insightful, it suffers from various limitations: First, only Germany as a target country and France as the only potential substitute is included. Second, unobserved country-specific effects are not controlled for. These effects are correlated if Germany and France share common cultural or reputational characteristics which make these countries more or less attractive for asylum seekers. The random-effects estimation, however, is based on the assumption that these effects are uncorrelated. Including country-fixed effects would be based on the much weaker assumption that the changes in these effects are uncorrelated, since correlation in the levels is removed (Plümper and Neumayer 2010). Third, to test whether France and Germany are regarded as substitutes by asylum seekers, this relationship could be modelled directly by estimating a spatial-lag model in which the asylum seekers in France are included as a control variable in the estimation model.

The analysis by Rotte et al. (1997) is extended by Vogler and Rotte (2000), which encompasses 86 Asian and African source countries to Germany. As a methodological advancement, they use country level fixed effects which allow controlling for unobserved, country-specific and time-invariant determinants of asylum outflows. The dependent variable is the migration rate, defined as the number of asylum seekers divided by the population of the source country. Among the factors that enhance the asylum migration are a high income differential between Germany and the source country, a higher GDP per capita in the source country, the intensity of political terror and the share of urban population. At a given level of income difference, a higher development level is associated with more asylum applications and there is some evidence that this effect is decreasing with rising living standards. The authors argue that this might be due to the availability of financial resources which are necessary to migrate. People from countries with low levels of GDP simply lack funds to make the journey to Germany and file an application there. With growing income, seeking asylum becomes less attractive. The factors with a negative impact on asylum migration are economic growth in the source country and bilateral trade. Separate estimations for Asian and African countries reveal a negative relationship between trade and asylum migration for Africa and a positive association for Asian sources. Increasing contacts which facilitate migration are provided as an explanation for the positive effect, while the negative correlation of trade and the political situation and its positive correlation with the numbers of asylum seekers explain the negative effect for Africa. Finally, the introduction of more restrictive German immigration laws in 1987 and 1993 deterred asylum seekers.

Exploiting a dyadic dataset which covers 20 source countries and 14 destination countries and spans the period from 1990 to 1999, Hatton (2004) estimates determinants of the share of asylum seekers in destination *i* at the total number of asylum applications from a source country *j*. In the focus of the analysis are potential spill-over effects from policy changes in other European destinations on the number of applicants in a given destination. Even though the author never names it, this is one of the very few works on spatial relationships in asylum migration. He finds that a higher GDP of a destination relative to the average of other potential destinations increases the share of applications in this destination. Again, a positive and highly significant effect of the past stock of nationals from the source in the host country is established. While a more restrictive asylum policy in a country reduces the share of asylum seekers in that country, no effect is found for the average asylum policy in other EU destinations. By adding the average of the asylum policy measure in other countries, Hatton (2004) implicitly estimates a so-called spatial-x model with a row-standardised unitary weighting matrix.[133] This gives equal weight to all other destinations, which might not be reasonable if some countries are closer substitutes for a given destination than others. Assessing the relative impact of changes in policy, GDP per capita and a general trend on the share of asylum seekers in each of the 14 destinations, the author finds that the country specific trend has the greatest influence, followed by GDP per capita and asylum policy. He argues that this is because all countries adopted more restrictive policies rather than due to the ineffectiveness of these policy measures.

The first comprehensive dyadic study is by Neumayer (2004) and examines what makes some Western European countries a more attractive destination for asylum seekers than others. Rather than explicitly controlling for source country characteristics, these effects are netted out by taking shares for each destination of applicants from each source as the dependent variable. The study covers 125 countries of origin, 17 destination countries and 18 years from 1982 to 1999. He finds that more asylum applications are lodged in richer countries, but neither a deterring effect of a higher unemployment rate nor of less social welfare spending. Contrary to expectations, strong economic growth in the destination country is associated with fewer

[133] In such a spatial-x model, one explanatory variable is spatially lagged and included in the estimation model.

asylum seekers. Regarding the political landscape of the target country, countries in which a right populist party gained a higher share in general elections are less attractive for migrants, whereas more come to countries with left-wing dominated countries. A higher recognition rate in the past also seems to encourage the influx of asylum seekers. Also, relatively fewer applications are filed in full Schengen member countries as opposed to non-member countries. The same is true for a larger past share of asylum applications. As expected, a higher share of asylum seekers comes from countries that are former colonies, speak the same language and are geographically closer. The negative impact of a lower recognition rate in the past and the lower relative attractiveness of full Schengen members suggests that destination countries are somewhat able to shift the burden of asylum seekers to other countries by setting a more restrictive asylum policy.

Working with a dyadic dataset that covers 40 source and 19 destination countries, Hatton (2009) analyses whether the widespread decline in the number of asylum applications since 2001 can be explained by tighter asylum policy. He finds that an increase in the stock of migrants from a given country by one percent increases the number of asylum application from the same source country by 0.33 percent. A one-percentage point increase in the unemployment rate in the destination country is associated with 10 percent fewer applications, while political terror as well as a more autocratic system in the source country increases the number of asylum seekers. Little evidence is found for a deterring effect of lower recognition rate. After 9/11, there were significantly fewer applications from Muslim countries; however, the effect disappears once an asylum policy measure is included. The recognition rate exhibits a certain degree of endogeneity if lower rates deter those most likely to be rejected. When instrumenting the recognition rate with changes in the policy towards manifestly unfounded claims and policy related to subsidiary status, the effect becomes positive and highly significant with a coefficient suggesting that a fall in the recognition rate by 10 percentage points is associated with 16 percent fewer applications. Adding controls for asylum policy, the effects of the policy components for access to the territory and processing of applications are negative as expected, but no evidence for a deterring effect of a tougher policy related to the welfare conditions of asylum seekers is found. The author finds that more restrictive policies concerning the access to the territory and the processing of applications account for 108,000 fewer asylum seekers per year or nearly one third of the overall decline in the application count from 2001 to 2006.

The dyadic part of Leblang's (2010) study covers asylum seeker flows from 125 sources to 20 destinations over 1983 to 2006. As in previous study, the author finds a positive effect of the bilateral migrant stock on the bilateral number of asylum applications. Similarly, a higher recognition rate and a larger income differential between the country of origin and the target country are associated with more asylum seekers. In line with Neumayer (2004), the government expenditures as a share of GDP seem not to have an impact; however, contrary to Neumayer (2004), no deterring effect of right populist parties can be established. As in Hatton (2009), a more restrictive asylum policy in the target country apparently discourages migrants to seek asylum in such a country. Regarding circumstances in the source nation, a violation of human rights and curtailed political rights increase the number of asylum seekers, whereas no statistically significant effect is found for ethnic wars and environmental crises. As in the monadic study, ODA as a share of GDP is negatively correlated with the number of asylum seekers. A decomposition of aid into post conflict, governmental and humanitarian aid reveals that the negative relationship only holds for the latter two but shows no effect of post conflict aid.

Summing up, while the effect of some variables is ambiguous and cross-study comparison is often hampered by different definition of variables, there is a clear pattern for several determinants whose influence is found in all or nearly all studies. Among these is the number of past asylum seekers or of official migrants, which indicates a network-effect among migrants as a higher number of fellow countrymen in the destination country reduces uncertainty and provides social amenities.[134] Another implication of these network effects is that it is much harder for governments to curtail the flow of asylum seekers once these networks are established, as Hatton (2004: 51) concludes: "Had policy been more pre-emptive then the channels of asylum migration would have become less entrenched and the subsequent

[134] This finding is corroborated by a non-quantitative analysis by Böcker and Havinga (1998) who find that over 90 percent of the asylum seekers from many origin countries are concentrated on less than five destination countries.

flows would have been easier to control." Other factors whose influence is unambiguous are political terror in the source country, which represents the original intent of granting asylum to refugees. Furthermore, a high GDP per capita differential between the source and the destination country fosters asylum migration, whereas most studies find a deterring effect of a more restrictive asylum policy in the host country. The same negative effect is found for geographical distance and high levels of unemployment in the target country. There is a plethora of other variables which have been tested, but these are either only included in one of the studies or the results are less consistent.

As argued in Section 3, there are several theoretical arguments why one would expect the numbers of asylum seekers from (to) one country to depend on the number of asylum seekers from (to) other source (destination) countries. This potential spatial dependence in asylum migration flows has been fully neglected so far. As discussed above, Rotte et al. (1997) and Hatton (2004) add third-country control variables to their estimation models, but do not refer to it as spatial dependence, which points into the direction that both are not aware that they are analysing spatial dependence. Consequently, they do not justify the choice of the implicitly used unitary weighting matrix.[135] Most importantly however, their approach does not provide insight on which countries are substitutes and complements from a migrant's point of view and whether asylum seekers are in fact deterred to other destination countries or whether an ever tighter asylum policy lowers the number of applicants in industrialised countries as a whole. The following empirical study seeks to answer this question. Furthermore, it is explored whether personal or human trafficker networks work across borders. Finally, the robustness of standard determinants of asylum migration to the inclusion of this spatial dependence is checked.

[135] Rotte et al. (1997) only include France as a third country in their Germany-specific case study. In this case, the choice of a weighting matrix becomes obsolete.

4.5. Data and methodology

Econometric techniques are used to test the theoretical predictions outlined above. Before the model and the estimation technique are addressed, the dependent variable and the control variables are presented and the calculation of the spatial lags is discussed in some detail.

The dependent variable

The dependent variable is the annual number of asylum applications lodged by nationals of source country *i* in target country *j*, i.e. the observational unit is a country-pair. Data is provided by UNHCR (2011), based on information reported by national governments of destination countries. The figures generally refer to the number of applicants or persons rather than the number of applications or families and only cover first-instance applications, i.e. they exclude repeat or appeal applications. Furthermore, only persons who officially lodged an application are considered. Hence, refugees who were not able or unwilled to apply for asylum are not covered (Neumayer 2005a).[136] Missing information is set to zero. However, some countries do not report detailed statistics for the whole period, for these, observations are left as missings before they start reporting asylum applications by country of origin. Following the example of Hatton (2009) and Leblang (2010), to reduce the skewness of the data and to mitigate the influence of large values, the log of the annual number of asylum seekers is taken.[137]

The dataset goes back until 1980 and covers the period including 2008, but the first two years are lost since some of the independent variables are lagged. The estimation data encompass 21 sources countries and up to 138 destination countries.[138] Tables 35 and 36 in the appendix provide a full list of source and target nations and also indicate the data availability for

---

[136] For instance, before 1991, only Europeans could apply for asylum in Italy. In general, this focus on refugees who successfully filed an asylum application introduces a sample selection bias: Only individuals who decided to migrate and to apply for asylum are included, but not those who stayed or became internally displaced persons. If those who leave share certain characteristics, the estimation sample is not representative for whole relevant population at risk. As a consequence, the results might not be generalisable, but are only valid for actual asylum seekers. Yet, data on persons thinking about leaving or refugees who did not reach their destination country to lodge an application is not available.

[137] The value of one was added to all values before taking the logarithm.

[138] 26 targets and 156 sources are used to calculate the spatial lags, but 5 and 18 are dropped because missing information for the control variables.

destination countries. The set of host countries comprises all major Western European countries, other large industrialised countries as well as three Eastern European countries which appeared on the destination list of asylum seekers only during the 1990s. Moore and Shellman (2007) criticise Neumayer (2004) for restricting the dataset to Western European destination countries as a source of severe sample selection bias and argue that all countries should be included as potential hosts. While this argument is certainly valid for their study on total refugees, the sample selection problem is less apparent for asylum seekers as these disproportionately often lodge their applications in industrialised countries.[139] In addition, asylum migration appears to follow different patterns than general refugee migration as with the exception of the United States and Germany, none of the top ten destination countries for refuges over the period 1965 to 1995 indentified by Moore and Shellman is one of the ten most popular targets for asylum seekers in 2005 to 2009. Furthermore, one of the main objectives of this work is to analyse whether tighter asylum policy deters asylum seekers to more lenient destinations and developed countries are notoriously active in trying to restrict the influx of applicants.[140] This is of particular interest, as developed countries do not only provide safety from political prosecution but also generally offer more social benefits which a are a burden to the taxpayer.

The independent variables

The set of source country characteristics encompasses:

- *Total population of source country* is used to control for the size of the source country and the pool of potential asylum seekers. Ceteris paribus, more asylum seekers should originate from larger countries.
- Since persons in working age are more likely to migrate, the *share of people between 15 and 64* at the total population is included as a control variable and a positive coefficient is expected.

[139] Between 2005 and 2009, on average 49.3 percent of all applications in the UNHCR (2011) database were filed in the 26 target countries, even though this group accounts for only 16.1 percent of all targets in the database. In any case, comprehensive data on asylum applications in developing countries is only available from late 1990s which would limit the study period considerably.

[140] Above all, important control variables are only available for a constrained set of industrialised nations.

- The *GDP per capita* level in constant 2005 USD is taken as an approximation for the development level and the economic situation in the source country. Escaping poverty and bleak economic prospects is one motivation to leave the home country. Little economic opportunities reduce the benefits of staying and should therefore increase the number of asylum seekers. However, migration is not for free and richer people can more easily raise the necessary funds for transport and traffickers. Hence, higher GDP per capita levels make migration more feasible, but less desirable (Vogler and Rotte 2000, Neumayer 2005a, Hatton 2009).[141]
- Devastating natural disasters can destroy livelihoods and social ties and therefore trigger emigration. The annual aggregate number of fatalities after droughts, earthquakes, epidemics, extreme temperatures, floods, earth slides, storms, volcano eruptions and wildfires is included (*natural disaster deaths*). A positive impact is anticipated.
- Also armed conflicts can cause individuals to flee their home county. As a measure for war intensity, the number of *battle fatalities* is included.[142] Data is taken from Lecina and Gleditsch (2005). If available, the best estimate of the number of people killed is used, otherwise the average of the lowest and highest estimate is taken. If more than one country is involved, the fatalities are equally allocated to territories. A positive sign is expected.
- The *political terror scale* index represents the scale of arbitrary imprisonment, torture, political murders and general violence. Data based on reports of the U.S. State department is used, but missing values replaced by the ratings based on Amnesty International reports. Further gaps are closed by inter- and extrapolation. The index ranges from 0 to 5 with higher values representing a worse situation. Since political repression is one of the original intents for providing asylum to refugees, a positive sign is expected.

---

[141] Unfortunately, reliable time-series data on income inequality is not available for a large set of developing countries. This would allow testing the argument made by Vogler and Rotte (2000) that the relative income situation matters (see Footnote 9).
[142] Natural disaster deaths and battle deaths enter the estimation in logs. To keep the zeros in the dataset, the value of one is added to all values, as with the dependent variable.

- *Democracy* is measured by the Polity2 variable taken from the PolityIV dataset (Marshall et al. 2006). The variable ranges from -10 (full autocracy) to 10 (full democracy). All else being equal, it is hypothesised that democratic systems impose less threat to the integrity of individuals and therefore cause lower numbers of refugees.
- *Urbanisation* represents the share of population living in urban areas. Since information is likely to better diffuse in urban areas, a positive sign is anticipated.

Controls for target country characteristics include:

- The *unemployment rate* in the target country, measured as unemployed people as a share of civilian labour force. A country with lower levels of unemployment should provide better work prospects which act as a pull-factor for immigrants. Furthermore, host countries with low unemployment rates should generally be in a better economic situation which enables them to extend the welfare state. Therefore, a negative sign is expected.
- The *population of the target country* rate is included as a measure for the size of the destination country. Larger countries have a higher absorptive capacity for immigrants and probably provide more entry points for unobserved illegal immigration. Furthermore, an asylum seeker is more likely to be aware of larger countries. Hence, a positive coefficient is expected.
- The *GDP per capita* of the destination country measured in constant 2005 USD is taken to control for the wealth of the host nation. Since richer countries should generally provide better employment opportunities and are able to spend more money for social purposes, it is expected that more migrants file for asylum in countries with a higher GDP per capita level.
- To control for changes in the domestic *asylum policy*, an index developed by Leblang (2010) is used. Starting from a value of zero in 1980 if no asylum policy is implemented in this year, the value of one is added for each year in which a policy is introduced that restricts asylum. In contrast, the value of one is deducted from the index in a given year for more liberal policy measures which makes asylum

immigration easier. The variable is lagged by one year to mitigate the apparent problem of endogeneity and to allow information to diffuse among potential asylum seekers. Since a tighter asylum policy should deter potential applicants, a negative sign is expected.[143]

Finally, the following two dyad-specific variables are included:

- As an additional measure for the generosity of a host country, the *recognition rate* is included in the estimation model. Data is provided by UNHCR (2011) broken down by country of origin and destination country. The theoretically correct measure, namely the percentage of recognised cases relative to the number of asylum claims lodged, is not available as many claims have not been decided in the period they were lodged and more detailed data is not available. Therefore, in line with UNHCR practice, the recognition rate is computed as the number of decisions recognising asylum claims in a given year as a share of the total number of claims decided upon.[144] The data are not without problems, however: First, they are not fully comparable over time (Vink and Meijerink 2003, Neumayer 2005b), as for some destination countries the data cover both first-instance and appeal decisions whereas for others only first-instance decisions are included. Second, in some host countries cases that are rejected on formal grounds enter the calculation, whereas they are excluded in other countries (Neumayer 2005b). Finally, as noted above, there is no information on when a decided application actually was lodged. It could be in current year, but also in any previous year and as a result, the recognition rate might not be a good proxy for the willingness to host asylum seekers of a destination country. For instance, this could be the case if there is a large backlog of cases after a mass influx of refugees as a consequence of a war in a major source country and these cases are rejected after several years because the political situation in

143 See footnote 122 for the prerequisites for an effectiveness of asylum policies. As discussed below, the estimations are based on dyad fixed effects, i.e. only variations over time are used to get estimates of the coefficients. Consequently, it is assumed that asylum seekers react to *changes* in asylum policy rather than the policy *level*.

144 The main measure covers only refugees recognized under the 1951 Geneva Convention. The results are, however, fully robust if also asylum seekers who are allowed to stay for other, mainly humanitarian, reasons are included.

the country of origin has stabilised in the meantime (Neumayer 2005b). While the first two problems are dealt with dyad fixed effects where variation across countries is neglected, the last issue is pure measurement error which attenuates the coefficient and makes statistically significant results less likely. As with the asylum policy measure, this variable is lagged to lessen the potential endogeneity and to allow time for the information to reach potential migrants.

- To test for the influence of official development assistance (ODA), the bilateral *ODA commitments* from the target to the source country are included. Since this is a mere control variable, contrary to Leblang (2010), different types of aid are not distinguished. The sign of the coefficient is not clear ex ante: One the one hand, more ODA could mitigate the causes of asylum migration and therefore reduce the number of asylum seekers. On the other hand, generous donors could enjoy a humanitarian reputation which attracts additional applicants in the hope of high recognition rates. Depending on which effect predominates, a negative or a positive sign emerges. However, both effects could also offset each other, rendering the effect close to zero and insignificant.

Unfortunately, time-invariant data on the stock of migrants from a source country in a given target country is not available to directly control for the personal migrant networks originating from other migrants from the same source country. As an approximation, the one-year lag of the dependent variable, i.e. the number of asylum seekers from source *i* in target *j* in year *t-1*, is added as a right-hand side variable. While this variable does not exactly cover the relevant measure, it has the advantage of not limiting the dataset apart from curtailing the time-dimension by one period. Furthermore, recent asylum migrants are more likely to be personally known by potential migrants and to act as a role model. In addition, their up-to-date experience might reduce the risk of migrating and help to make contacts to trustworthy human traffickers. In addition, path-dependency could also stem from sunk costs from establishing migration routes and networks incurred which incurred to people smugglers and agents and which they are reluctant to give up (Pierson 2000).

Spatial lags

To analyse whether the number of asylum seekers from (to) a given source (target) country, depends on the number of asylum seekers in other source (target) countries, a spatial-y model is estimated. In such a model, the dependent variable for other observations is included as a right-hand side variable (Anselin 1988). The connectivity between countries is represented by a weighting matrix, in which each cell indicates the relationship either between two source countries or two target countries. The dependent variable is then multiplied with the weighting matrix to obtain the spatially lagged variable. Such a spatial lag resembles a temporal lag, but rather than representing values of the dependent variable from a former period in time, it takes the values of the dependent variable from geographically remote observations.[145]

The weighting matrices for specific source contagion are derived from the theoretical discussion in Section 3. Unfortunately, modelling human trafficking routes is not straightforward, as they are often unknown, complex and the popularity of single routes has risen and fallen over time, often in respond to new anti-smuggling measures taken by target country governments (UNODC 2011). However, since both theoretical arguments stipulate a positive spatial dependence in the number of asylum seekers between geographically close source countries, they are unlikely to offset each other. Hence, while a positive spatial lag is in line with both arguments, the two effects cannot be disentangled in the empirical operationalisation due to a lack of detailed data on the actual migrant routes between a source and a target country. Specifically, the following set of weighting matrices is tested:

- *Contiguity*: The first weighting matrix is a simple binary matrix, in which a cell takes the value of one if two source countries share a common border and zero otherwise. In all weighting matrices, the cross-diagonal elements are set to missing as a source country cannot spatially depend on itself. This weighting matrix allocates the same weight to each neighbouring country and captures both the idea of transnational migration network effects and increased efficiency of human smuggling networks.

[145] Most of the weighting matrices used here are based on some sort of geographical distance; however, the concept of distance is also transferable to other forms of similarity between different countries.

Since adjacent countries often share cultural similarities, a higher number of asylum seekers from neighbouring source countries in a specific target country could reduce the risk of migration for asylum seekers from a given source country and therefore increase their numbers. By the same token, asylum seekers from neighbouring countries could gain access to well established migration routes.

- The first weighting matrix assumes that all adjacent countries exert influence on a specific source country, but no other country that does not share a border with this given source country. This assumption might be too restrictive if also asylum seekers from not-adjacent countries, but from the same region provide a positive network benefit or if migration routes connect a whole region to major destination countries. Therefore, the second weighting matrix for specific source contagion is *Common region*, which is again a binary matrix allocating equal weight to all other source countries from the same region. According to the World Bank classification, five regions are distinguished: Latin America and the Caribbean, Sub Saharan Africa, Northern Africa and the Middle East, South Asia and East Asia and the Pacific.
- While the second weighting matrix is broader than the first, it still assumes that the influence of other countries of origin abruptly ends at a region's border. The third weighting matrix, *Inverse distance*, allows every other source country in the sample to impact the number of asylum seekers from a given source nation. Since cultural similarity and access to migration networks are bound to decrease with distance, each cell contains the inverse of the distance between two source countries. It is unlikely, however, that both effects operate over a very large distance. Thus, inverse distance is calculated as $1/\text{distance}^2$ to allow the effect to decay quickly with distance. The distance measure taken is the population weighted distance between two countries provided by CEPII and described in Mayer and Zignago (2006).
- Finally, personal migration networks which facilitate transition to the host country might not be constrained to cultural similarity which stems from geographical proximity of the source country but also originate from a shared language. Hence, a larger community of migrants who speak the same language lowers the cost of asylum migration from source countries where this language is spoken. The weighting matrix *Common language* gives equal weight to all other source countries which share a

common language with a given source country. Since not the official tongue but the actually spoken language matters, a dyad is classified as sharing a common language if it is spoken by at least nine percent of the population in both countries.[146] Data is taken from the same source as distance.

In its most parsimonious cross-sectional form, abstracting from the temporally lagged dependent variable and the control variables, the estimation model for specific source contagion reads as follows (Neumayer and Plümper 2010):

$$y_{ij} = \rho \sum_{k \neq i} w_{ik} y_{kj} + \varepsilon_{ij} \qquad (15)$$

The dependent variable y_{ij} is the number of asylum seekers from source i in target j, y_{kj} is the number of asylum seekers from other source countries k to the same target j and w_{ik} represents the weighting matrix which models the connectivity between source country i and other sources k. The spatial lag parameter ρ is to be estimated and indicates, if statistically significant, that the number of asylum seekers of the dyad ij depends on the number of asylum seekers from other sources to the same destination.

The main theoretical argument for specific target contagion is that asylum seekers are deterred by a restrictive asylum policy in the destination country. If potential migrants are not completely discouraged by a tighter policy stance in their preferred target country but directed to other destination countries which are substitutes from the migrant's point of view, the number of asylum seekers in one destination country should spatially depend on the number of applications in other targets. Since asylum policy is the main theoretical driver of such a diversion, the policy index provided by Leblang (2010), which is also included as a control variable, is used as a measure for the restrictiveness of a target's asylum policy. If the theoretical argument is correct, asylum seekers should be directed to countries with a more lenient asylum policy and attracted from countries with a more restrictive policy, while

[146] This measure might not fully reflect network effects from source countries in which a large number of languages is spoken. Since asylum seekers are not distinguished by their mother tongue, a more detailed analysis is not feasible. The results are not changed if the weighting matrix is based on official languages.

countries with the same policy should not influence the number of applications in the target country under observation. In principal, these differences in asylum policy could be expressed by the difference in the asylum policy measure between target *j* and target *m*; however, negative values are not feasible in a row-standardised weighting matrix as positive and negative values would offset each other. Therefore, two different weighting matrices are created: the first one measures the difference in asylum policy to more lenient countries, the second one the difference in asylum policy to more restrictive countries. Negative values are set to zero and both spatial lags are estimated simultaneously in one model. Using these weighting matrices allocates a higher weight to countries in which the asylum policy differs greatly from the target in question and no weight to countries with the same policy restrictiveness. Specifically, the following set of weighting matrices is used:[147]

- *Difference in asylum policy (more restrictive targets)*: As outlined, this weighting matrix allocates a higher weight to targets with a more restrictive asylum policy than target *j* and no weight to countries with a more lenient or the same policy. A negative sign is expected as lower numbers of applications in target *m* should lead to a higher number of applications in target *j* if asylum seekers are attracted by the relatively more liberal policy in target *j*.
- *Difference in asylum policy (less restrictive targets)*: In contract, this weighting matrix allocates a higher weight to targets with a less restrictive asylum policy than target *j* and no weight to countries with a more restrictive or the same policy. Again, a negative sign is expected as higher numbers of applications in target *m* should lead to a lower number of applications in target *j* if asylum seekers are directed to target *m* by the relatively more liberal policy in target *m*.
- *Difference in asylum policy (more restrictive targets) x Common language*: Here, the weighting matrix is multiplied with a dummy taking the value of one if the same language is spoken in target *j* and target *m*. As a consequence, no weight is given to countries with the same asylum policy and to countries with a different national language, while countries with the same language receive weight according to their

[147] Contrary to specific source contagion, the spatial lags with the contiguity and common region are not tested. The reason is that the main target region, the European Union, is very heterogeneous. Contiguity is dropped since geographical proximity and not the existence of a common border should influence the degree of substitutability.

relative policy restrictiveness.[148] This weighting matrix is used to test whether countries with the same language are closer substitutes than countries with a different language. The rationale behind this is that if a migrant has some knowledge of a foreign language, assimilation in a country in which this language is spoken is easier than in a country with a different national language. If the preferred host country's asylum policy is too restrictive, the asylum seeker might regard other destinations with the same language as alternatives. As before, a negative sign is expected, but the size of the coefficient of the spatial lag should be larger if the hypothesis, that same language countries are closer substitutes, is true.

- *Difference in asylum policy (less restrictive targets) x Common language*: This weighting matrix is simply the counterpart of the previous matrix. Again and for the same reason, a negative effect is anticipated.
- *Difference in asylum policy (more restrictive targets) x Inverse distance*: Here, the weighting matrix is multiplied with the inverse distance ($1/\text{distance}^2$) between target *j* and target *m*. Here, no weight is allocated to countries with the same asylum policy, but the influence of targets with a more restrictive policy increases with their relative restrictiveness and decreases with distance. This matrix is used to test the hypothesis that geographically closer countries are closer substitutes than more remote targets.[149] A negative sign which is larger than the sign of the spatial lag with the non-interacted weighting matrix is anticipated.
- *Difference in asylum policy (less restrictive targets) x Inverse distance*: Again the counterpart of the previous weighting matrix to model a decreasing influence over space of countries with a more liberal policy. The expected sign is negative.

Formally, specific target contagion can be expressed as follows (Neumayer and Plümper 2010):

[148] It is noteworthy that this is not an interaction term in the sense of Neumayer and Plümper (2012) to measure whether the spatial effect differs systematically across countries. They suggest an interaction of the spatial lag with a dummy accounting for the differences between targets. Here, the not *the effect of the spatial lag* varies across countries, but the *influence* of other targets.

[149] Here, not the distance between source country *i* and target *j* is meant, but the distance between target *j* and other targets *m*.

$$y_{ij} = \rho \sum_{m \neq j} w_{jm} y_{im} + \varepsilon_{ij} \qquad (16)$$

The number of asylum seekers in country j originating from country i (y_{ij}) is explained by the weighted number of asylum applications from the same source country j in other destinations m. Here, the weighting matrix w_{jm} models the connectivity between the targets j and m.

The weighting matrices for all spatial lags are row-standardised. As a consequence, the spatial lags represent the weighted average number of asylum seekers in other sources (specific source contagion) and in other targets (specific target contagion), respectively, rather than the weighted total number of asylum seekers in other countries of origin and destinations.[150] While this row-standardisation is common practice in the applied empirical literature on spatial dependence, it needs to be theoretically justified as it changes the relative importance of impacting countries and hence the estimation results (Plümper and Neumayer 2010). Using relative rather than absolute weights removes all level effects in the weighting matrix; for instance, taking the simple binary contiguity weighting matrix discussed above, the total effect of spatial dependence is assumed to be the same for Portugal which just has Spain as a neighbouring country as for Germany with its nine neighbouring target nations in the sample (Neumayer and Plümper 2012). Row-standardisation has the additional virtue that the coefficients can be interpreted as elasticities since the dependent variable and the spatial lag are both expressed in the same unit, namely log of the number of asylum seekers.

Tables 30 and 31 provide summary statistics and data sources for all variables and the spatial lags.

[150] To row-standardise a weighting matrix, each cell in a row of the matrix is divided by its row sum. As a consequence, the weights add up to one for each row.

Table 30: Summary statistics and data sources of the dependent and the control variables

Variable	Mean	Std. Dev.	Min.	Max.	Source
Dependent Variable					
Asylum seekers (log)	1.719	2.218	0	11.656	UNHCR (2011)
Source county controls					
Total population (ln)	16.000	1.541	12.357	21.004	World Bank (2011)
Share of people 15-64	58.652	6.725	46.196	82.719	World Bank (2011)
Share of urban population	47.132	22.852	4.480	100	World Bank (2011)
GDP per capita (n)	7.239	1.330	4.308	11.043	UNCTAD (2011)
Natural disaster deaths (ln)	2.106	2.502	0	12.612	EM-DAT (2011)
Battle deaths (log)	1.458	2.867	0	11.796	Lecina and Gleditsch (2005)
Political terror scale	2.775	1.087	1	5	Wood and Gibney (2010)
Democracy	0.552	6.810	-10	10	Marshall et al. (2006)
Target country controls					
Unemployment rate	7.645	3.928	0.185	24.172	OECD (2011)
Total population (ln)	16.738	1.087	15.082	19.534	World Bank (2011)
GDP per capita (ln)	10.192	0.511	8.406	11.116	UNCTAD (2011)
Asylum policy (t-1)	-0.618	2.513	-9	7	Leblang (2010)
Dyad-specific controls					
Recognition rate (t-1)	0.166	0.272	0	1	UNHCR (2011)
ODA commitments (ln, t-1)	6.893	7.563	0	22.732	Nielson et al. (2010)

Note: With the exception of the recognition rate (26,061), the summary statistics are given for the most comprehensive estimation model with 62,752 observations.

Table 31: Summary statistics of the spatial lags

Variable	Mean	Std. Dev.	Min.	Max.
Specific source contagion				
W: Contiguity	1.675	1.790	0	9.811
W: Common region	1.469	1.326	0	7.565
W: Inverse distance	1.409	1.330	0	8.929
W: Common language	1.239	1.393	0	10.955
Specific target contagion				
W: Asylum policy difference (more restrictive targets)	2.064	1.987	0	10.463
W: Asylum policy difference (less restrictive targets)	1.321	1.433	0	9.177
W: Asylum policy difference (more restrictive targets) x Common language	1.052	1.928	0	11.208
W: Asylum policy difference (less restrictive targets) x Common language	0.0889	1.686	0	10.205
W: Asylum policy difference (more restrictive targets) x Inverse distance	2.130	2.123	0	10.463
W: Asylum policy difference (less restrictive targets) x Inverse distance	1.434	1.696	0	9.378

Notes: W: denotes the weighting matrix used; All spatial lags are derived from own calculations; The summary statistics are given for the most comprehensive estimation model with 62,752 observations for specific source contagion and 62,736 observations for specific target contagion.

Model and methodology

In the most general form, the full estimation model reads as follows:

$$y_{ijt} = \alpha_{ij} + \beta y_{ij(t-1)} + \rho' X_{ijt} + \pi' W_{ijt} + \delta' I_i \cdot T_t + \mu' J_j \cdot T_t + \varepsilon_{ijt} \quad (17)$$

Here, y_{ijt} is the number of asylum seekers from source country i in target country j in year t, while $y_{ij(t-1)}$ is the one-year lag of the dependent variable. The constant α_{ij} represents a dyad-specific fixed effect which captures all time-invariant characteristics which vary across dyads, but not within a country-pair. The matrix X_{ijt} contains the spatial lags described above, whereas all source-specific, target-specific and dyad-specific control variables are represented by W_{ijt}. Finally, $I_i \cdot T_t$ and $J_j \cdot T_t$ are time trends for the source country i and the target country j, respectively.

The dependent variable, the number of asylum seekers, is a classical example of a count variable. Since the sample variance greatly exceeds the sample mean, a negative binominal model would be the most appropriate estimation technique. Here, equation (4) is nevertheless estimated with ordinary least squares (OLS), mainly because of two reasons: One the one hand, while spatial dependence is well understood in the workhorse OLS, it is not obvious whether the estimation of a spatial lag model is easily transferable to estimation techniques for count data. On the other hand, and more importantly, the negative binominal model fails to yield results either due to infeasible initial values or since a flat region in the estimation is discovered. Using OLS in count data has one minor drawback: the predictions made from an OLS model can be negative, which is not reasonable for asylum seekers. However, this problem is avoided by taking the log of the dependent variable and making predictions is not the main objective of this analysis in any case. Estimating a dynamic model such as the one in equation (4) with a fixed effects model also introduces a Nickell (1981) bias; yet this bias diminishes as the number of periods T gets large and the dataset covers the period from 1981 to 2008. In such a dynamic model, the coefficients estimated only represent the short-term effect. To calculate the long-term effect, the coefficient of the temporal lag of the dependent variable has to be taken into account.

The main challenge when estimating a spatial lag model is to establish a causal relationship between the spatially lagged dependent variable and the dependent variable. A statistically significant coefficient of the spatial lag does not necessarily indicate such causality, but could solely represent spatial clustering or unobserved spatial heterogeneity. Spatial clustering denotes a situation where close countries share certain observable characteristics, whereas unobserved spatial heterogeneity means that these factors either cannot be observed or appropriate measures are not available.[151] In the context of asylum migration, such spatial clustering could occur if geographically close countries share common characteristics or impacts, e.g. political unrest in the whole region, a drought affecting more than one country, which lead - together with similar bilateral characteristics with the destination country - to an simultaneous increase of asylum seekers of different source countries from a region to the same host. Unobserved spatial heterogeneity occurs for example if the measures used provide a poor approximation of the asylum emigration causing phenomenon and the real effect is partly subsumed in the error term. If this is true for two or more geographically close countries, there are spatial patterns in the error term. Other factors, such as a cultural rootedness and the willingness to abandon the home country are unobservable and might be similar across adjacent countries. Similarly, if geographically close target countries (e.g., Scandinavian countries) share certain unobserved characteristics which are attractive for asylum seekers such as the reputation of being a lenient destination country, a simultaneous, yet independent, rise in the number of asylum seekers in more than one host country could be the consequence. Spatial clustering in target countries occurs if a group of countries, e.g. Mediterranean EU members, have a tight asylum policy and low recognition rate.

Following the suggestions of Plümper and Neumayer (2010), several steps in the model specification are taken to mitigate the problem of unobserved spatial heterogeneity, to account for spatial clustering and to make measuring spatial contagion more likely: First, the model specification is as broad as possible to control for a wide range of source-, target-, and dyad-specific factors which could influence the number of asylum seekers in a dyad to account for spatial clustering. This reduces the risk of omitted variable bias in the spatial lags if these determinants are correlated with the spatial lags. Of particular importance are source country specific control variables which account for influences that could trigger a simultaneous increase in the number of asylum seekers in adjacent countries, such as those controlling for natural disasters or conflicts which might affect more than one source country. Second, and most important, the estimation is based on dyad fixed-effects. While this approach has the

[151] Close countries in this context are mostly, but not necessarily, geographically close countries. Also countries with a similar historical background such as Commonwealth Membership might share certain observable or unobservable features, even if they are thousands of miles apart.

drawback that no estimates for important time-invariant bilateral characteristics such as distance and common language can be obtained, all observable and unobservable time-invariant factors are automatically controlled for as they are subsumed in the dyad-specific constant.[152] The estimation is based on over time variation of the variables only, whereas all level effects are removed. For the spatial lag, this has the advantage that also all spatial clustering and unobserved spatial heterogeneity in levels is eliminated. As a consequence, not high *numbers* of asylum seekers from adjacent countries in the same destination country lead to a positive coefficient of the spatial lag, but simultaneous *increases* (or decreases) in the number of asylum seekers from close source countries in the same host country, which is a much stronger prerequisite. Third, the inclusion of the temporally lagged dependent variable also captures common trends and additionally accounts for temporal dynamics (Plümper and Neumayer 2010). Finally and in addition to the recommendations of Plümper and Neumayer (2010), a source and destination country time trend is added, which controls for any time trend in the number of asylum seekers that is specific to a country of origin and a target country.[153] These dyad-partner time effects account for fundamental differences in the variation of the number of asylum emigrants and immigrants across countries over time. If these differences are not fully captured by the control variables and are correlated with the spatial lags, these variables of interest could suffer from an omitted variable bias. They also effectively control for general trends in the size of asylum migration, for instance, if over time more people have learnt about the possibility to seek asylum abroad, transportation has become cheaper or simply if changes in reporting have occurred.

[152] The inclusion of dyad-fixed effects leads to a small sample-bias known as Nickell-Hurwicz-bias (Franzese and Hays 2007). With more than 26,000 and 62,000 observations this should not impose a problem in the present estimation results.

[153] Note that this procedure differs from Chapter 2, in which only a global time trend is added, mainly for two reasons: First, bilateral aid was expressed in shares rather than in absolute figures, which removes very much of the time trend and mitigates the influence of unobserved country heterogeneity. Second, in contrast to migration decisions, the aid allocation decision is taken by government authorities rather than individuals, which limits heterogeneity and allows controlling more directly for influencing factors.

4.6. Main results and robustness checks

The presentation of the results is organised as follows: First, the results for specific source and specific target contagion are shown individually. Since the effects are not mutually exclusive, they are subsequently estimated together.

Table 32 displays the estimation results for specific source contagion. Since data on the recognition rate is limited for earlier years and completely unavailable for some dyads, the inclusion considerably restricts the sample size: while the number of dyads decreases from 2,893 to 2,418, the number of observation drops from nearly 63,000 to slightly more than 26,000. This sample reduction is likely to be non-random; therefore, the estimation results are presented for two specifications: In Model I to Model VI, the recognition rate is not controlled for, whereas this variable is added in Model VII to Model XII. When interpreting the results, one has to keep in mind that the estimation is based on dyad fixed effects, that is, all unobservable (and theoretically observable) country-pair specific time invariant influences are automatically accounted for. Such estimation is based on variation over time only. It is obvious that the inclusion of source and target specific time trends picks up a lot of the variation over time and that it might render control variables, which have an up- or downward trend or little variation over time, insignificant. As expected, many of the control variables are insignificant. The population of the target country and partly GDP per capita of the target country are dropped because of collinearity, which is not surprising since both variables have a fairly constant growth rate over time and are linearised by taking the log. For the few statistically significant control variables, the effect size and even the sign strongly depends on the model specification. To check the plausibility of the controls, the estimations without the source and target specific time trend, but with a *t-1* set of year dummies are presented in Table 37 in the Appendix. Such a global trend accounts for general trends in the number of asylum seekers, but it assumes that the trend is the same for each dyad. Formally, the estimated model can be written as follows:

$$y_{ijt} = \alpha_{ij} + \beta y_{ij(t-1)} + \rho' X_{ijt} + \pi' W_{ijt} + \delta' T_t + \varepsilon_{ijt} \tag{18}$$

where T_t represents the set of year dummies to control for a global trend.

The control variables are now very much in line with expectations. First of all, the one-year lag of the dependent variable is positive and highly significant in all specifications, providing evidence for a high path-dependency in the number of asylum applications. A one percent increase in the past number of asylum seekers is associated with a more than 0.6 percent increase in the contemporary application level. As discussed above, while this clearly provides evidence for the migration network effects, it might also be due to persistence in the migration routes. More asylum seekers originate from larger source countries while the number decreases with rising GDP per capita, while higher levels of unemployment in the target country discourage migrants from lodging their application in this country. There is only limited evidence that a higher share of people aged 15 to 64 is associated with higher migration outflows, however, a higher degree of urbanisation in the source country leads to more asylum migrants. As argued by Neumayer (2005a), this could be due to a better diffusion of information in urban areas. An intense armed conflict, measured by the number of battle fatalities, triggers migration outflows, whereas no such an effect can be detected for natural disasters. The original intent of asylum is to provide protection to people who are prosecuted in their home country. In line with this definition and with expectations, higher values on the political terror scale lead to more refugees: a one point increase on a scale ranging from 0 to 5 is associated with a 11 to 14 percent higher number of applications in the short-run, depending on model specification. Even after controlling for the threat of arbitrary imprisonment, torture and political murder, significantly fewer asylum applications come from democratic systems. A one point increase on the -10 to 10 democracy scale leads to a decrease in the number of refugees by 0.8 to 1.6 percent in the short-run. The asylum policy measure provided by Leblang (2010) has the expected sign, as a more restrictive asylum policy is associated with lower application numbers. A one-point increase in the index, which runs from -10 to 10, leads however to a mere 1.2 to 2.9 percent short-run decrease in the number of asylum seekers. Either, this little size of the effect could be due to a lack of knowledge of asylum seekers about the actual asylum policies in potential destination countries, or asylum policy is no determinant of the choice where to lodge the asylum application. This latter point might be because refugees are desperate about leaving their country and therefore are not deterred by tighter asylum policy or because they are optimistic about being accepted despite a toughening of asylum policy. Furthermore, the time span from the application to the final decision and an eventual deportation is regularly several years. An applicant might hope for a more lenient asylum procedure in the future or that the situation in his or her home country has improved by the time his application is denied. Finally, the variable is only a very crude measure of asylum policy, as it gives equal weight to all policy changes, no matter how significant they are.

The recognition rate, introduced in Model VII to XII, is positive and highly statistically significant in all specifications. The coefficient indicates that a 10 percentage point higher recognition rate in the previous period increases contemporary asylum applications in the short-run by around 1.3 percent. To calculate the long-run effect, the coefficient of the temporally lagged dependent variable has to be taken into account. Specifically, the coefficient of a given control variable is divided by one minus the coefficient of the one-year lag of the dependent variable (Egger and Merlo 2007). For Model VII, for instance, the long-run effect of a ten percentage point increase in the recognition rate is 3.7 percent. The negative impact of lower recognition rates could be explained one the one hand by deterrence, i.e. potential asylum seekers decide not to migrate or to lodge their application in another country. On the other hand, as noted by Hatton (2004), migrants may also decide not to apply for asylum and remain underground rather than risking rejection and removal.[154] The latter point might be particular relevant for those with weak asylum claims and in destinations with a flourishing underground economy. Some of the variables react very strongly to the introduction of the recognition variable and to the considerably curtailed data set in Models VII to XII.[155] The population of the target country is positive and highly significant in the more comprehensive data indicating that larger countries experience more applications. At the same time, the coefficient of the population in source countries more than triples with a much more precise estimate despite a lower number of observations. In contrast, the ODA commitment variable loses significance and the positive link between ODA and asylum seekers can only be

[154] The same argument applies to a tighter asylum policy regime.

[155] Running the estimation models in the new data sample without the recognition rate variable reveals that all changes are in fact due to the smaller sample size rather than caused by the inclusion of the new variable (results not shown).

established in the larger sample. This points into the direction that the reputation effect of bilateral ODA outweighs the mitigation of the causes of asylum effect.

Table 32: OLS estimation results for specific source contagion with source and target specific time trend

Model	I	II	III	IV	V	VI	VII	VIII	IX	X
Asylum seekers (ln, t-1)	0.529***	0.532***	0.528***	0.538***	0.525***	0.531***	0.530***	0.532***	0.542***	0.524***
	(69.37)	(69.10)	(68.66)	(72.44)	(66.83)	(45.41)	(44.40)	(44.88)	(46.52)	(43.81)
W: Contiguity (t-1)	0.0766***				0.0391***	0.0963***				0.0457*
	(9.18)				(3.50)	(6.64)				(2.22)
W: Common region (t-1)		0.120***			0.0632***		0.189***			0.150***
		(7.57)			(3.52)		(7.59)			(5.24)
W: Inverse distance (t-1)			0.102***		0.0432**			0.116***		0.0387
			(8.44)		(2.64)			(6.07)		(1.54)
W: Common language (t-1)				0.0489***	0.00885				0.0326*	-0.0271
				(4.03)	(0.69)				(2.00)	(-1.52)
Recognition rate (t-1)						0.0599	0.0628	0.0602	0.0609	0.0616
						(1.67)	(1.73)	(1.67)	(1.69)	(1.70)
Unemployment rate target	0.0391***	0.0199*	-0.0621***	-0.0405***	-0.0795***	-0.122*	-0.117*	-0.121*	-0.118*	-0.120*
	(4.44)	(2.54)	(-5.91)	(-4.99)	(-6.88)	(-2.44)	(-2.38)	(-2.41)	(-2.36)	(-2.43)
Population target (ln)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
GDP per capita target (ln)	3.509***	1.896***	-0.00788	1.116***	-1.275***	n/a	n/a	n/a	n/a	n/a
	(5.53)	(9.20)	(-0.03)	(4.34)	(-4.49)					
Asylum policy target (t-1)	-0.000418	0.00287	0.0638***	0.0523***	-0.0752***	0.264*	0.263**	0.265**	0.265**	0.263**
	(-0.03)	(0.24)	(4.97)	(4.31)	(-5.39)	(2.54)	(2.63)	(2.60)	(2.59)	(2.61)
Population source (ln)	-311.3	-268.9	-458.4	-554.7	-160.9	2165	2133	1695	1730	2250
	(-0.31)	(-0.45)	(-0.40)	(-0.44)	(-0.37)	(0.40)	(0.40)	(0.32)	(0.32)	(0.42)
GDP per capita source (ln)	22.06	-115.4	72.35	-323	-57.83	-663	-652.9	-519.7	-530	-688.8
	(0.31)	(-0.45)	(0.40)	(-0.44)	(-0.36)	(-0.41)	(-0.40)	(-0.32)	(-0.32)	(-0.42)

Table 32: OLS estimation results for specific source contagion with source and target specific time trend (continued)

Model	I	II	III	IV	V	VI	VII	VIII	IX	X
Share urban population source	-2.371	8.512	9.152	6.716	-4.445	-2.322	-2.293	-1.827	-1.861	-2.417
	(-0.31)	(0.45)	(0.41)	(0.44)	(-0.36)	(-0.41)	(-0.40)	(-0.32)	(-0.33)	(-0.42)
Share population 15-64 source	-4.486	9.412	-6.538	12.78	-4.305	83.07	81.84	65.05	66.35	86.33
	(-0.31)	(0.45)	(-0.40)	(0.44)	(-0.36)	(0.40)	(0.40)	(0.32)	(0.32)	(0.42)
Disaster deaths source (ln)	1.958	-9.574	4.853	-9.108	4.516	1.323	1.307	1.009	1.032	1.384
	(0.32)	(-0.44)	(0.42)	(-0.44)	(0.36)	(0.37)	(0.37)	(0.29)	(0.29)	(0.39)
Battle fatalities source (ln)	1.815	27.88	-3.491	5.801	-0.712	22.34	22.01	17.54	17.89	23.21
	(0.32)	(0.45)	(-0.41)	(0.44)	(-0.35)	(0.41)	(0.40)	(0.32)	(0.32)	(0.42)
Political terror source	-20.52	-34.54	6.365	-45.82	-53.56	-331	-326	-259.4	-264.6	-343.9
	(-0.31)	(-0.44)	(0.44)	(-0.44)	(-0.36)	(-0.41)	(-0.40)	(-0.32)	(-0.32)	(-0.42)
Democracy source	-3.747	0.538	0.658	-11.79	0.236	-100.1	-98.65	-78.49	-80.06	-104
	(-0.31)	(0.45)	(0.44)	(-0.45)	(0.26)	(-0.41)	(-0.40)	(-0.32)	(-0.32)	(-0.42)
ODA commitment (t-1)	0.00247**	0.00243*	0.00259**	0.00264**	0.00239*	0.00135	0.00122	0.00142	0.00149	0.0012
	(2.62)	(2.57)	(2.75)	(2.80)	(2.53)	(0.85)	(0.76)	(0.90)	(0.94)	(0.75)
Constant	5023	4185	6653	10014	3436	-34654	-34179	-27162	-27694	-36018
	(0.31)	(0.45)	(0.40)	(0.44)	(0.37)	(-0.40)	(-0.40)	(-0.32)	(-0.32)	(-0.42)
Observations	62,752	62,752	62,752	62,752	62,752	26,061	26,061	26,061	26,061	26,061
Number of dyads	2,893	2,893	2,893	2,893	2,893	2,418	2,418	2,418	2,418	2,418
R-squared	0.642	0.642	0.642	0.641	0.642	0.682	0.683	0.682	0.681	0.683

Notes: W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on source and target specific time trends not shown; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

Turning to the spatial lags in the original model specification (Table 32) it can be seen that they are positive as theoretically predicted and statistically significant at least at the five percent level as long as they are estimated individually. A ten percent increase in the average number of asylum seekers from neighbouring countries in a given destination country in the last year is associated with a 0.77 percent increase in applications from a given source in that destination country in the short-run (Model I). The corresponding long-term effect is 1.63 percent. For the average number of asylum seekers from the same source region, the effect is somewhat larger (Model II), as it is if other source countries are weighted according to their inverse distance to the source country in question (Model III). The short-term effect of a ten percent increase is only 0.49 percent for asylum seekers from other source countries in which the same language is spoken (Model IV). Since these effects are not mutually exclusive and to assess their relative importance, all four spatial lags are estimated simultaneously in Model V. Since the spatial lags are positively correlated as they measure similar concepts, it is not surprising that the coefficients become smaller. The spatial lag with the common language weighting matrix becomes insignificant, while the other three remain positive and significant. This indicates that geographical proximity rather than a shared language matters for cross-border network effects. The result also is in line with the hypothesis that scale effects in migration routes facilitate asylum migration, as these are effects should decrease with distance. In the remaining models in which the recognition rate is additionally controlled for, with the exception of the spatial lag with the common language weighting matrix, the coefficients of the spatial lags are larger. If all spatial lags are again estimated together (Model X), only the spatial lags with the contiguity and the common region weighting matrix remain significant. This, however, is a consequence of the considerably reduced sample size rather than of adding the recognition rate as a control variable (results not shown) and indicates that missing data for the recognition rate is non-random and introduces a sample selection bias.

A brief look at the spatial lags in the model specification with the global time trend (Table 37) shows, that their inclusion hardly affects the coefficient size and the significance level of the remaining control variables. This provides confidence in the results of previous studies, which do not control for the potential spatial dependence, as there is no evidence that they suffer from an omitted variable bias. The results corroborate the previous findings, even though the coefficients of the spatial lags are generally somewhat smaller than in the main specification, yet positive and statistically significant throughout. However, they lose their significance if they are estimated simultaneously in the smaller dataset which includes the control variable for the recognition rate (Mode XII), but this again is an artefact of the smaller sample size.

For specific target contagion where the number of asylum seekers from source *i* in target *j* depends on the number of applicants from source *i* in other targets *m*, theory predicts a negative effect. As can be seen from the results in Table 33, this relationship is confirmed in the empirical analysis. To save space, only the coefficients of the temporal lag, the spatial lags and the recognition rate variable are shown. The estimation models include source and target specific time trends to account for unobserved spatial heterogeneity. As before, the temporal lag is positive and highly significant in all specifications. Model I presents the estimation results for the spatial lag where weights are allocated according to the difference in asylum policy, but these weights do not depend on other factors. Both spatial lags, for more restrictive and more liberal targets, are negative and significant at the 0.1 percent level. The coefficients indicate that in the short-run, a ten percent decrease in the number of applications from a source country in other more restrictive targets are associated with a 1.1 percent increase in the target under observation. Similarly, a ten percent increase in the number of applicants from a country of origin in more liberal destination countries leads to a decrease of 1.9 percent in a given target country. The corresponding long-term effects are 2.5 percent and 4.0 percent, respectively. This provides evidence that asylum seekers are deterred by restrictive asylum policy and attracted by a more lenient policy and that domestic asylum policy creates spill-over effects onto all other destination countries.

Table 33: OLS estimation results for specific target contagion with source and target specific time trend

Model	I	II	III	IV	V	VI
Asylum seekers (ln, t-1)	0.527***	0.541***	0.539***	0.533***	0.542***	0.543***
	(71.31)	(73.80)	(73.39)	(45.98)	(47.36)	(47.03)
W: Asylum policy difference (more restrictive targets)	-0.111***			-0.113***		
	(-13.91)			(-9.77)		
W: Asylum policy difference (less restrictive targets)	-0.187***			-0.162***		
	(-18.31)			(-8.90)		
W: Asylum policy difference (more restrictive targets) x Common language		0.00315			-0.0412***	
		(0.49)			(-4.51)	
W: Asylum policy difference (less restrictive targets) x Common language		-0.0172**			-0.0627***	
		(-2.74)			(-7.22)	
W: Asylum policy difference (more restrictive targets) x Inverse distance			-0.0518***			-0.0607***
			(-7.51)			(-6.36)
W: Asylum policy difference (less restrictive targets) x Inverse distance			-0.0609***			-0.0489***
			(-8.35)			(-4.24)
Recognition rate (t-1)				0.0635	0.0675	0.057
				(1.78)	(1.88)	(1.59)
Observations	62,763	62,763	62,763	26,061	26,061	26,061
Number of dyads	2,893	2,893	2,893	2,418	2,418	2,418
R-squared	0.648	0.641	0.643	0.685	0.682	0.682

Notes: Set of control variables as in Table 32; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on source and target specific time trends not shown; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

If the influence of more liberal and more restrictive destinations is limited to those with a common language (Model II), the spatial lag for more restrictive targets becomes positive, but far from significant. However, the spatial lag for more liberal targets is still negative and significant, albeit with a short-run elasticity of 0.017 it is considerably smaller than the effect in the first specification. While this result is still in line with the previous finding, no evidence is found for the hypothesis that targets with the same language are closer substitutes. Finally, if the influence of other targets is modelled not only to depend on the difference in asylum policy, but also to decrease with distance between the targets, the spatial lags are negative and highly significant (Model III). Yet, the coefficient size again is smaller rather than larger than in the first specification, indicating that geographically closer countries are not necessarily closer substitutes. This result is somewhat at odds with the anecdotal evidence presented by other authors (see Section 3). A possible explanation is that closer countries have a more similar asylum policy and therefore do exert less influence on each other. On the one hand, the difference in asylum policy might increase with distance between targets. On the other hand, the weighting matrix in Model III allocates less influence to more distant countries. If the former effect outweighs the latter, the total distance effect is positive. Model IV to VI presents the estimation results in the smaller sample, after controlling for the recognition rate. This variable has the right sign, but is marginally insignificant as a consequence of the target specific time trend. The spatial lags are comparable to the first sample, but both spatial lags, where the effect is restricted to targets with a common language, are negative and significant (Model V). The relative effect sizes of the spatial lags, however, are as before.

So far, specific source and specific target contagion are estimated separately; however, there is no reason to assume that the two effects are independent or mutually exclusive. Therefore, Table 34 depicts the results of estimation models which include both types of spatial dependence. Due to space restrictions, only the non-interacted weighting matrix for difference in asylum policies is taken as a measure for specific target contagion and added to the previous set of estimation models. Again, all estimations encompass a source and target specific time trend to account for unobserved spatial heterogeneity. As before, all spatial lags are in line with expectations: the coefficients for the spatial lags controlling for specific source contagion are positive and highly significant, whereas the spatial lags for specific target contagion are negative and significant at the 0.1 percent level throughout. Yet, the coefficients are affected by estimating the spatial effects simultaneously. Most importantly, the elasticity of the temporal lag increases from slightly more than 0.5 to around 0.75, indicating a much higher time dependency than before. As a consequence, the difference between the short-term and the long-term effect becomes larger. Second, the coefficients of the spatial lags that measure specific source contagion are also inflated, with an increase between 29 percent for spatial lag with the common region weighting matrix (Model II) and 84 percent for the spatial lag with the common language weighting matrix (Model IV). This pattern persists if all spatial lags are estimated together (Model V) and with the exception of the spatial lag using the common region weighting matrix also in the smaller dataset, but after controlling for the recognition rate (Model VI to X). There are also some changes in the significance levels in the most comprehensive Models V and X, since the spatial lag using the common language weighting matrix remains significant in Model V, as does the spatial lag with the inverse distance weighting matrix in Model X. Third, the spatial lag for the difference in asylum policy for more liberal targets becomes considerably smaller, while the spatial lag for targets with a more restrictive asylum policy is hardly affected.

Table 34: OLS estimation results for specific source and specific target contagion with source and target specific time trend

Model	I	II	III	IV	V	VI	VII	VIII	IX	X
Asylum seekers (ln, t-1)	0.743***	0.750***	0.737***	0.763***	0.730***	0.799***	0.805***	0.797***	0.819***	0.789***
	(119.4)	(126.7)	(118.4)	(137.5)	(114.3)	(110.8)	(111.2)	(108.9)	(123.6)	(104.0)
Specific source contagion										
W: Contiguity (t-1)	0.110***				0.0466***	0.108***				0.0505***
	(16.88)				(5.74)	(13.27)				(4.52)
W: Common region (t-1)		0.155***			0.0638***		0.144***			0.0669***
		(16.73)			(5.91)		(11.64)			(4.65)
W: Inverse distance (t-1)			0.160***		0.0780***			0.140***		0.0599***
			(17.65)		(6.74)			(13.10)		(4.16)
W: Common language (t-1)				0.0898***	0.0287***				0.0679***	0.0131
				(11.49)	(3.46)				(7.66)	(1.38)
Specific target contagion										
W: Asylum policy difference (more restrictive targets)	-0.107***	-0.106***	-0.107***	-0.109***	-0.106***	-0.097***	-0.094***	-0.098***	-0.099***	-0.094***
	(-17.19)	(-16.72)	(-17.39)	(-17.66)	(-16.78)	(-10.79)	(-10.30)	(-10.93)	(-11.11)	(-10.31)
W: Asylum policy difference (less restrictive targets)	-0.139***	-0.142***	-0.138***	-0.139***	-0.143***	-0.111***	-0.116***	-0.110***	-0.113***	-0.115***
	(-16.15)	(-16.28)	(-16.12)	(-16.09)	(-16.38)	(-8.50)	(-8.92)	(-8.46)	(-8.73)	(-8.76)
Recognition rate (t-1)						0.0138	0.0158	0.0212	0.00821	0.0255
						(0.47)	(0.53)	(0.72)	(0.28)	(0.86)
Observations	62,752	62,752	62,752	62,752	62,752	26,061	26,061	26,061	26,061	26,061
Number of dyads	2,893	2,893	2,893	2,893	2,893	2,418	2,418	2,418	2,418	2,418
R-squared	0.876	0.876	0.877	0.875	0.877	0.883	0.883	0.883	0.882	0.884

Notes: Set of control variables as in Table 32; W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on source and target specific time trends not shown; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

4.7. Conclusion

The aim of this chapter was to analyse whether the number of asylum seekers from one source country in a host nation depends on the number of asylum migrants from other source countries in the same destination and on the number of applications in other target countries. There are two main arguments for the existence of the first phenomenon, called specific source contagion: First, personal migrant networks might not be constrained to migrants from the same source country, but they also work across borders, just as ethnicities are not geared to national boundaries. These personal networks facilitate transition and reduce assimilation and risk cost in the host country by supplying information about agents and assistance in the search for accommodation and employment. Thereby, existing migrants provide a positive externality which makes asylum migration from geographically close source countries more likely. Second, most asylum seekers rely on the services of people smugglers who invest time and money to gather information about the asylum regime in host countries and the most promising routes as well as to make contacts with carriers. Migration along well established routes reduces both costs and risk for the migrant or makes migration possible in the first place. This is not only true for refugees on the main routes, but also for those stemming from countries through which these main routes run and those using smaller feeder routes to access the major routes. The second form of spatial dependence, namely specific target contagion, arises if asylum seekers are deflected by a more restrictive asylum policy in one target country and incentivised to lodge their application in another destination country. Many policy measures, for example visa restrictions or lower recognition rates, are targeted against asylum seekers from a specific source country. As a consequence, the number of asylum seekers from a source country in a given target country should depend on the number of applications from that country of origin in other target countries.

The results confirm the existence of both specific source and specific target contagion. As expected, the number of asylum seekers from a source country increases the size of the migration flow from geographically close other source countries. Only limited evidence, however, is found that these network effects extend to migrants from other remote source countries in which the same language is spoken as in the source country in question. The estimation results also confirm the existence of specific target contagion, since an increase in the number of asylum seekers in all other destination countries is associated with a decrease in the number of applications in a given target country. This lends support to the deflection hypothesis.

Both findings have important policy implications: Hatton (2004) argues that the time-dependency limits the effectiveness of asylum policies to reduce the number of asylum seekers. This is because the stock of migrants who are already in the host country decreases the costs for subsequent migrants, but those already residing in the target country are not affected by the policy measures. The same argument applies for policy actions that are targeted against refugees from a given source country. Targeted asylum policy falls short since it does not affect the number of applications from other source countries, but these migrants provide a benefit to migrants from the country against which the policy measures are directed and thereby undermine their effectiveness. Despite some efforts to harmonise asylum policy to share the burden of asylum, particularly in the European Union, most asylum measures remain unilateral. Policy restrictions provide a negative externality to other target countries since the flow of asylum seekers is deflected and these other countries experience higher application numbers. As a consequence of such a beggar-thy-neighbour policy, governments have an incentive to engage in a race-to-the-bottom and lower the standards until they run into severe conflicts with a nation's obligation under human rights treaties such as the Geneva Convention (Neumayer 2004). This negative externality provides a further reason for policy harmonisation and common standards in asylum policy, so that not only a minimum treatment to refugees is guaranteed, but also that the negative impact of a government's actions on other countries is limited.

References

Aldridge, Fiona and Sue Waddington (2001), *Asylum Seekers' Skills and Qualifications Audit Pilot Project*, Leicester: National Institute for Adult Continuing Education.

Anselin, Luc (1988), *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer.

Beine, Michael, Frédéric Docquier, and Caglar Özdem (2011), *Diaspora Effects in International Migration – Key Questions and Methodological Issues*, World Bank Policy Research Working Paper 5721, Washington, DC: World Bank.

Bilsborrow, Richard and Hania Zlotnik (1994), The Systems Approach and the Measurement of the Determinants of International Migration, in: von der Erf, Rob and Liesbeth Heering (eds.), *Causes of International Migration*, Brussels: Statistical Office of the European Communities, pp. 61-76.

Bloch, Alice (2002), *Refugees' opportunities and barriers to employment and training*, Department of Work and Pensions Report No. 179, London: HMSO.

Böcker, Anita and Tetty Havinga (1998), Asylum Applications in the European Union: Patterns and Trends and the Effects of Policy Measures, *Journal of Refugee Studies*, 11(3), p. 245-266.

Borjas, George (1994), The Economics of Immigration, *Journal of Economic Literature*, 32(4), pp. 1667-1717.

Brochmann, Grete (1995), *Bosnian Refugees in the Nordic Countries. Three Routes to Protection*, Nordic Comparative Studies on the Reception of Refugees Series, Oslo: Ministry of Local Government and Labour.

Brücker, Herbert and Philip J.H. Schröder (2010), *Migration Regulation Contagion*, Working Paper 10-22, Aarhus: Aarhus University, Department of Economics.

Carey-Wood, Jenny (1995), *The settlement of refugees in Britain*, Home Office Research Study No. 141, London: HMSO.

Cho, Seo-Young, Axel Dreher and Eric Neumayer (2011), *The Spread of Anti-Trafficking Policies – Evidence from a New Index*, CESIFO Working Paper No. 3376, Munich: CESIFO.

Davenport, C.A., Moore, W.H. and Poe, S.C. (2003), Sometimes you just have to leave: domestic threats and forced migration, 1964–1989, *International Interactions*, 29, pp. 27–55.

Egger, Peter and Valeria Merlo (2007), The Impact of Bilateral Investment Treaties on FDI Dynamics, *The World Economy*, 10, pp. 1536-1549.

EM-DAT (2011), *The International Disaster Database*, available at: http://www.emdat.be/database, last access: 22. April 2011.

Facchini, Giovanni and Anna-Maria Mayda (2008), From Individual Attitudes Towards Migrants to Migrant Outcomes, *Economic Policy*, 56, pp. 651-713.

Franzese, Robert and Jude Hays (2007), Empirical models in international capital-tax competition, in: Read, C. and G.N. Gregiriou (eds.), *International taxation handbook, Policy, practice, standards and regulation*, Amsterdam: Elsevier.

Gibney, Matthew J. and Randall Hansen (2005), Asylum Policy in the West: Past Trends, Future Possibilites, in: Borjas, George J. and Jeff Crisp (eds.), *Poverty, International Migration and Asylum*, Hampshire: Palgrave Macmillan, pp. 70-96.

Hailbronner, Kay (2000), *Immigration and Asylum Law and Policy of the European Union*, The Hague: Kluwer Law International.

Hansen, Randall and Desmond King (2000), Illiberalism and the New Politics of Asylum: Liberalism's Dark Side, *Political Quarterly*, 71(4), pp. 396-403.

Hartog, Joop and Aslan Zorlu (2009), How important is homeland education for refugees' economic position in the Netherlands?, *Journal of Population Economics*, 21(3), pp. 679-684.

Hatton, Timothy (2004), Seeking Asylum in Europe, *Economic Policy*, 19(38), pp. 5-62.

Hatton, Timothy (2009), The Rise and Fall of Asylum: What Happened and Why?, *The Economic Journal*, 119, pp. 183-213.

Hatton, Timothy and Jeffrey Williamson (2005), What fundamentals drive world migration?, in: Borjas, George and Jeff Crisp (eds.), *Poverty, International Migration and Asylum*, Basingstoke: Palgrave Macmillan, pp. 15-38.

Havinga, Tetty and Anita Böcker (1999), Country of asylum by choice or by chance: asylum-seekers in Belgium, the Netherlands and the UK, *Journal of Ethnic and Migration Studies*, 25(1), pp. 43-61.

Holzer, Thomas, Gerald Schneider, and Thomas Widmer (2000), The Impact of Legislative Deterrence Measures on the Number of Asylum Applications in Switzerland (1986-1995), *International Migration Review*, 34(4), pp. 1182-1216.

Husted, Leif, Helena Skyt Nielsen, Michael Rosholm, Nina Smith (2000), *Employment and Wage Assimilation of Male First Generation Immigrants in Denmark*, IZA Discussion Paper 101, Bonn: IZA.

International Organization for Migration (2000), *Migrant Trafficking and Human Smuggling in Europe*, Geneva: United Nations.

ISSP (1995), *The International Social Survey Programme*, 1995 Survey, available at: http://zacat.gesis.org/webview/, last access: 22. August 2011.

ISSP (2003), *The International Social Survey Programme*, 2003 Survey, available at: http://zacat.gesis.org/webview/, last access: 22. August 2011.

Krahn, Harvey, Tracey Derwing, Marlene Mulder and Lori Wilkinson (2000), Educated and Underemployed: Refugee Integration into the Canadian Labour Market, *Journal of International Migration and Integration*, 1(1), pp. 59-84.

Leblang, David (2010), *The Supply of Foreign Aid and the Demand for Asylum*, Charlottesville, VA: University of Virginia.

Lecina, Bethany and Nils Petter Gleditsch (2005), Monitoring Trends in Global Combat: A New Dataset of Battle Deaths, *European Journal of Population*, 21(2-3), pp. 145-116.

Marshall, Monty, Keith Jaggers and Ted Gurr (2006), *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2009*, available at: http://www.systemicpeace.org/polity/polity4.htm, last access: 23. June 2011.

Martin, Susan, Andrew I. Schoenholtz and David Fisher (2005), The Impact of Asylum on Receiving Countries, in: Borjas, George J. and Jeff Crisp (eds.), *Poverty, International Migration and Asylum*, Hampshire: Palgrave Macmillan, pp. 99-120.

Massey, Douglas (1990), The Social and Economic Origins of Immigration, *Annals of the American Academy of Political and Social Science*, 510, pp. 60-70.

Mayer, Thierry and Soledad Zignago (2006), Notes on CEPII's distances measures, available at: http://www.cepii.fr/distance/noticedist_en.pdf.

Moore, Will H. and Stephen M. Shellman (2004), Fear of Prosecution: Forced Migration 1952-1995, *Journal of Conflict Resolution*, 40(5), pp. 723-745.

Moore, Will H. and Stephen M. Shellman (2006), Refugee or Internally Displaced Person? To Where Should One Flee?, *Comparative Political Studies*, 39(5), pp. 599-622.

Moore, Will H. and Stephen M. Shellman (2007), Whither Will They Go? A Global Study of Refugee's Destinations, 1965-1995, *International Studies Quarterly*, 51, pp. 811-834.

Neumayer, Eric (2004), Asylum Destination Choice: What Makes Some West European Countries More Attractive Than Others?, *European Union Politics*, 5(2), pp. 155-180.

Neumayer, Eric (2005a), Bogus Refugees? The Determinants of Asylum Migration to Western Europe, *International Studies Quarterly*, 49, pp. 389-409.

Neumayer, Eric (2005b), Asylum Recognition Rates in Western Europe – Their Determinants, Variation and Lack of Convergence, *Journal of Conflict Resolution*, 49(1), pp. 43-66.

Neumayer, Eric (2006), Unequal Access to Foreign Spaces: How States Use Visa Restrictions to Regulate Mobility in a Globalised World, *Transactions of the British Institute of Geographers*, 31(1), pp. 72-84.

Neumayer, Eric and Thomas Plümper (2010), Spatial Effects in Dyadic Data, *International Organization*, 64(1), pp. 145-166.

Neumayer, Eric and Thomas Plümper (2012), Conditional Spatial Policy Dependence: Theory and Model Specification, *Comparative Political Studies*, forthcoming.

Nickell, Stephen (1981), Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49, pp. 1417-1426.

Nielson, Daniel L., Ryan M. Powers, and Michael J. Tierney (2010), *Broad Trends in Foreign Aid: Insights from PLAID 1.9.1.*, Conference paper presented at Aid Transparency and Development Finance: Lessons and Insights from AidData, Oxford, U.K. 22-25 March 2010.

Noll, Gregor (2000), *Negotiating Asylum*, The Hague: Martinus Nijhoff Publishers.

OECD (2011), *Labour force statistics for OECD member countries*, available at: http://stats.oecd.org, last access: 22. April 2011.

Olson, Mancur (1965), *The Logic of Collective Action*, Cambridge, MA: Harvard University Press.

Petros, Melanie (2005), *The costs of human smuggling and trafficking*, Global Migration Perspectives No. 31, Geneva: Global Commission on International Migration.

Pierson, Paul (2000), Increasing Returns, Path Dependence, and the Study of Politics, *American Political Science Review*, June, pp. 251-266.

Plümper, Thomas and Eric Neumayer (2010), Model Specification in the Analysis of Spatial Dependence, *European Journal of Political Research*, 49(3), pp. 418-442.

Roberts, Adam (1998), More Refugees, Less Asylum: A Regime in Transformation, *Journal of Refugee Studies*, 11(4), pp. 375-395.

Robinson, Vaughan and Jeremy Segrott (2002), *Understanding the decision-making of asylum seekers*, Home Office Research Study 243, London: Home Office Research.

Rotte, Ralph, Michael Vogler, and Klaus Zimmermann (1997), South-North Refugee Migration: Lessons for Development Cooperation, *Review of Development Economics*, 1(1), pp. 99-115.

Schilling, S.R. (1995), Refugees and Immigration in Europe and the Third World, in: Heckmann, F. And Bosswick, W. (eds.), *Migration Policy: A Comparative Perspective*, Stuttgart: Enke, pp. 263-266.

Schmeidl, Susanne (1997), Exploring the Causes of Forced Migration: A Pooled Time-Series Analysis, 1971-1990, *Social Science Quarterly*, 78(2), pp. 284-308.

Schuster, Liza (2000), A Comparative Analysis of the Asylum Policy of Seven European Governments, *Journal of Refugee Studies*, 13(1), pp. 118-132.

Suhrke, Astri (1998), Burden-sharing during Refugee Emergencies: The Logic of Collective Action versus National Action, *Journal of Refugee Studies*, 11(4), pp. 316-415.

Teitelbaum, Michael (1984), Immigration, Refugees and Foreign Policy, *International Organization*, 38(3), pp. 429-450.

The International Herald Tribune (2011), *A harsh life for some who land in Europe; Looking for wealth, many African refugees instead find dead end*, 14. May 2011.

Thielemann, Eiko (2004), Why Asylum Policy Harmonisation Undermines Refugee Burden-Sharing, *European Journal of Migration and Law*, 6, p. 47-65.

Thielemann, Eiko (2006a), The Effectiveness of Governments' Attempts to Control Unwanted Migration, in Parsons C. and Smeeding T. (eds.), *Immigration and the Transformation of Europe*, Cambridge: Cambridge University Press, pp.442-472

Thielemann, Eiko (2006b), *Burden Sharing: The International Politics of Refugee Protection*, The Center of Comparative Immigration Studies Working Paper No. 134, San Diego: University of California.

Thielemann, Eiko, Richard Williams and Christina Boswell (2010), *What system of burden-sharing between Member States for the reception of asylum seekers?,* PE 419.620, Brussels: European Parliament.

UNCTAD (2011), UNCTADStat, available at: http://unctadstat.unctad.org, last access: 01. May 2011.

UNHCR (2010), *UNHCR urges EU and FRONTEX to ensure access to asylum procedures, amid sharp drop in arrivals via the Mediterranean*, Briefing Note, 10. December 2010, available at: http://www.unhcr.org/4d022a946.html, last access: 20. August 2011.

UNHCR (2011), *Statistical Online Population Database*, available at: www.unhcr.org/statistics/populationdatabase, last access: 23. April 2011.

UNODC (2011), *The role of organized crime in the smuggling of migrants from West Africa to the European Union*, Vienna: United Nations Office on Drugs and Crime.

Van Moppes, David (2006), *The African Migration Movement: Routes to Europe*, Working Papers Migration and Development Series, Report No. 5, Nijmegen: Radboud University.

Vink, Maarten and Frits Meijerink (2003), Asylum Applications and Recognition Rates in EU Member States 1982-2001: A Quantitative Analysis, *Journal of Refugee Studies*, 16(3), pp. 297-315.

Vogler, Michael and Ralph Rotte (2000), The effects of development on migration: Theoretical issues and new empirical evidence, *Journal of Population Economics*, 13, pp. 485-508.

Wood, Reed M. and Mark Gibney (2010), The Political Terror Scale (PTS): A Re-introduction and a Comparison to CIRI, *Human Rights Quarterly*, 32(2), pp. 367-400.

World Bank (2011), *World Development Indicators*, available at: http://databank.worldbank.org, last access: 23. April 2011.

Zetter, Roger, David Griffiths, Silva Ferretti, and Martyn Pearl (2003), *An assessment of the impact of asylum policies in Europe 1990-2000*, Home Office Research Study 259, London: Home Office Research.

Appendix

Table 35: List of target countries and first year of detailed data availability

Australia (1996), Austria (1980), Belgium (1980), Canada (1989), Czech Republic (1990), Denmark (1984), *Finland (1981)*, France (1981), Germany (1980), Greece (1980), Hungary (1995), *Iceland (1997)*, Ireland (1991), Italy (1980), *Japan (1996)*, *Luxembourg (1995)*, Netherlands (1980), *New Zealand (2001)*, Norway (1985), Poland (1991), Portugal (1982), Spain (1982), Sweden (1981), Switzerland (1980), United Kingdom (1980), United States (1987)

Notes: List of target countries used to calculate the spatial lags; due to data availability of the control variables, countries in *italics* are excluded from the estimation.

Table 36: List of source countries

Afghanistan, Albania, Algeria, Angola, *Antigua and Barbuda*, Argentina, Armenia, Azerbaijan, Bahrain, Bangladesh, *Barbados*, Belarus, *Belize*, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, *Brunei Darussalam*, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, *Cape Verde*, Central African Republic, Chad, Chile, China, Colombia, Comoros, Democratic Republic of the Congo, Republic of the Congo, Costa Rica, Cote d'Ivoire, Croatia, Cuba, Czech Republic, Djibouti, *Dominica*, Dominican Republic, Ecuador, Arab Republic of Egypt, El Salvador, Equatorial Guinea, Eritrea, Estonia, Ethiopia, Fiji, Former Yugoslavia (Serbia, Montenegro, Kosovo), Gabon, The Gambia, Georgia, Ghana, *Grenada*, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, *Hong Kong*, Hungary, India, Indonesia, Islamic Republic of Iran, Iraq, Israel, Jamaica, Jordan, Kazakhstan, Kenya, Democratic Republic of Korea, Republic of Korea, Kuwait, Kyrgyz Republic, Lao People's Democratic Republic, Latvia, Lebanon, Lesotho, Liberia, Libya, Lithuania, *Macao*, Macedonia, Madagascar, Malawi, Malaysia, *Maldives*, Mali, Mauritania, Mauritius, Mexico, Moldova, Mongolia, Morocco, Mozambique, Myanmar, Namibia, Nepal, Nicaragua, Niger, Nigeria, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Qatar, Romania, Russian Federation, Rwanda, *Samoa*, *Sao Tome and Principe*, Saudi Arabia, Senegal, *Seychelles*, Sierra Leone, Singapore, Slovak Republic, Slovenia, Solomon Islands, Somalia, South Africa, Sri Lanka, *St. Kitts and Nevis*, *St. Lucia*, *St. Vincent and the Grenadines*, Sudan, *Suriname*, Swaziland, Syrian Arab Republic, Tajikistan, Tanzania, Thailand, Timor-Leste, Togo, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Uganda, Ukraine, United Arab Emirates, Uruguay, Uzbekistan, Venezuela, Vietnam, *West Bank and Gaza,* Republic of Yemen, Zambia, Zimbabwe

Notes: List of sources countries used to calculate the spatial lags; due to data availability of the control variables, countries in *italics* are excluded from the estimation.

Table 37: OLS estimation results for specific source contagion with general time trend

Model	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Asylum seekers (ln, t-1)	0.628***	0.609***	0.618***	0.609***	0.623***	0.608***	0.659***	0.647***	0.649***	0.647***	0.653***	0.644***
	(92.30)	(78.90)	(78.81)	(78.68)	(86.70)	(75.60)	(63.95)	(58.99)	(57.82)	(58.62)	(62.26)	(57.27)
W: Contiguity (t-1)		0.0682***				0.0530***		0.0536***				0.0197
		(9.00)				(5.04)		(4.32)				(0.99)
W: Common region (t-1)			0.0511***			-0.0165			0.0691***			0.028
			(4.42)			(-1.00)			(3.87)			(1.14)
W: Inverse distance (t-1)				0.0792***		0.0393*				0.0650***		0.026
				(8.03)		(2.48)				(4.47)		(1.08)
W: Common language (t-1)					0.0356***	-0.00393					0.0463***	0.0107
					(3.38)	(-0.30)					(3.42)	(0.65)
Recognition rate (t-1)							0.125***	0.125***	0.127***	0.126***	0.127***	0.127***
							(3.63)	(3.63)	(3.69)	(3.66)	(3.70)	(3.68)
Unemployment rate target	-0.008***	-0.0075**	-0.0076**	-0.0073**	-0.0076**	-0.0074**	-0.020***	-0.021***	-0.021***	-0.021***	-0.021***	-0.021***
	(-3.34)	(-3.24)	(-3.27)	(-3.14)	(-3.27)	(-3.19)	(-3.62)	(-3.80)	(-3.82)	(-3.71)	(-3.74)	(-3.83)
Population target (ln)	0.977***	0.740**	0.783**	0.704**	0.866***	0.732**	0.641	0.525	0.461	0.5	0.559	0.45
	(4.14)	(3.13)	(3.28)	(2.97)	(3.63)	(3.07)	(1.27)	(1.04)	(0.91)	(0.99)	(1.11)	(0.89)
GDP per capita target (ln)	0.332***	0.283***	0.299***	0.292***	0.312***	0.287***	0.193	0.152	0.144	0.162	0.164	0.139
	(4.49)	(3.79)	(4.00)	(3.90)	(4.21)	(3.83)	(0.82)	(0.64)	(0.61)	(0.69)	(0.70)	(0.59)
Asylum policy target (t-1)	-0.012***	-0.014***	-0.013***	-0.014***	-0.013***	-0.014***	-0.028***	-0.029***	-0.029***	-0.029***	-0.029***	-0.029***
	(-3.61)	(-3.99)	(-3.89)	(-4.09)	(-3.77)	(-4.05)	(-4.47)	(-4.52)	(-4.52)	(-4.53)	(-4.56)	(-4.56)
Population source (ln)	0.228**	0.201**	0.222**	0.192**	0.213**	0.193**	0.687***	0.657***	0.730***	0.665***	0.697***	0.687***
	(3.19)	(2.76)	(3.05)	(2.64)	(2.93)	(2.65)	(4.84)	(4.59)	(5.13)	(4.66)	(4.91)	(4.82)
GDP per capita source (ln)	-0.214***	-0.223***	-0.219***	-0.217***	-0.215***	-0.221***	-0.341***	-0.349***	-0.341***	-0.340***	-0.339***	-0.343***
	(-8.18)	(-8.60)	(-8.25)	(-8.27)	(-8.13)	(-8.50)	(-6.29)	(-6.42)	(-6.22)	(-6.21)	(-6.20)	(-6.28)

Table 37: OLS estimation results for specific source contagion with general time trend (continued)

Model	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Share urban population source	0.0099***	0.0103***	0.0103***	0.0107***	0.0100***	0.0104***	0.0179***	0.0179***	0.0176***	0.0180***	0.0177***	0.0177***
	(5.05)	(5.28)	(5.26)	(5.52)	(5.11)	(5.42)	(3.71)	(3.69)	(3.59)	(3.73)	(3.65)	(3.64)
Share population 15-64 source	0.00405	0.00547	0.00553	0.00700*	0.00365	0.00618*	0.00732	0.00744	0.00947	0.00924	0.00637	0.00879
	(1.40)	(1.90)	(1.90)	(2.42)	(1.26)	(2.10)	(1.32)	(1.34)	(1.70)	(1.65)	(1.15)	(1.55)
Disaster deaths source (ln)	0.00142	0.00147	0.0014	0.00111	0.00146	0.00131	-0.00515	-0.00512	-0.00535	-0.00536	-0.00529	-0.00533
	(0.68)	(0.71)	(0.67)	(0.53)	(0.70)	(0.63)	(-1.60)	(-1.59)	(-1.66)	(-1.66)	(-1.65)	(-1.66)
Battle fatalities source (ln)	0.020***	0.019***	0.019***	0.019***	0.019***	0.019***	0.024***	0.023***	0.024***	0.024***	0.024***	0.023***
	(7.30)	(7.12)	(7.21)	(7.03)	(7.18)	(7.06)	(6.02)	(5.84)	(5.92)	(5.87)	(5.96)	(5.84)
Political terror source	0.114***	0.114***	0.115***	0.113***	0.114***	0.113***	0.137***	0.140***	0.140***	0.139***	0.138***	0.140***
	(15.35)	(15.43)	(15.47)	(15.31)	(15.40)	(15.33)	(10.82)	(11.02)	(10.97)	(10.98)	(10.86)	(11.00)
Democracy source	-0.008***	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***	-0.01***	-0.016***	-0.016***	-0.016***	-0.015***	-0.016***
	(-5.21)	(-5.66)	(-5.60)	(-5.80)	(-5.46)	(-5.71)	(-4.54)	(-4.77)	(-4.80)	(-4.78)	(-4.78)	(-4.87)
ODA commitment (t-1)	0.00231*	0.00216*	0.00233*	0.00227*	0.00238*	0.00216*	0.000237	0.000267	0.000246	0.000283	0.000155	0.000251
	(2.35)	(2.21)	(2.37)	(2.33)	(2.43)	(2.22)	(0.15)	(0.17)	(0.16)	(0.18)	(0.10)	(0.16)
Constant	-22.26***	-17.45***	-18.69***	-16.95***	-19.97***	-17.30***	-22.45*	-19.59*	-19.77*	-19.56*	-20.91*	-18.80*
	(-5.16)	(-4.01)	(-4.23)	(-3.88)	(-4.55)	(-3.91)	(-2.35)	(-2.05)	(-2.07)	(-2.04)	(-2.20)	(-1.96)
Year fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	62,752	62,752	62,752	62,752	62,752	62,752	26,061	26,061	26,061	26,061	26,061	26,061
Number of dyads	2,893	2,893	2,893	2,893	2,893	2,893	2,418	2,418	2,418	2,418	2,418	2,418
R-squared	0.51	0.512	0.511	0.511	0.51	0.512	0.482	0.483	0.482	0.483	0.482	0.483

Notes: W: denotes the weighting matrix used for calculating the spatial lags; Robust standard errors clustered on dyads used; t-values shown in parentheses; Includes dyad-level fixed effects; Coefficients on (t-1) year dummies not displayed; * statistically significant at 0.05, ** 0.01, or *** 0.001 level.

5. Conclusion

Contrary to the work on spatial dependence in monadic data, empirical studies on spatial dependence in dyadic datasets are still scarce. So far, only a very limited number of studies have been published, but all of them show that spatial dependence is a wide-spread phenomenon also in dyadic data. This thesis contributed to filling this gap in the literature by analysing spatial dependence in the diffusion of double taxation treaties (DTT), official development assistance (ODA) and asylum migration. In all three applications, comprehensive evidence for spatial dependence is found. In the first case, it is shown that the propensity that a given country-pair signs a tax treaty is positively influenced by treaty-signing behaviour of other countries. Particularly, a treaty conclusion will be more likely if competitors for foreign direct investment entered such a DTT. This result confirms the theoretical argument by Baistrocchi (2008) that net-capital-importers find themselves in a situation which can be described as a prisoners' dilemma: while net-capital-importers individually should not sign tax treaties with major capital exporters, collectively they have an incentive to do so. This is because on the one hand a tax treaty with a major capital exporter could lead to a considerable loss in tax revenues for net-capital-importers. On the other hand, such a treaty can provide a competitive advantage in the rivalry for foreign capital. If one country enters a DTT with an important source country of foreign investment, it gains an advantage over other countries without such a treaty. Therefore, other focal countries have an incentive to follow suit and to negotiate such a treaty themselves. However, in such a situation, no country has a competitive advantage, but all net-capital-importers face decreasing tax revenues. In the second case, it is demonstrated that aid from a donor to a recipient is not independent from aid from other donors to the same beneficiary. Yet, there is little evidence for strategic interaction among donors, in the course of which aid is used as a mean to pursue their national economic and military interests. Rather, the results suggest that smaller donors tend to follow the aid allocation decisions of the important players. The third application analysed spatial dependence in bilateral asylum flows, both on the source and the target side. It was argued that personal migrant networks which work across borders reduce the risk of transition and therefore foster asylum migration. Furthermore, the design and increasing efficiency of international human trafficking routes lower the cost of transportation from countries which are adjacent to major source countries. Finally, it has been demonstrated that asylum seekers are deflected by a tighter asylum policy in one country and encouraged to lodge their application in a target country with a more liberal stance towards asylum seekers.

All three examples show that spatial dependence in dyadic data is a widespread phenomenon. Practically all empirical studies which are based on a dyadic framework, for instance those employing the standard gravity-model, can be augmented to test for spatial contagion. This is appropriate if theory suggests the existence of such spatial dependence among dyads, dyad-members, sources or targets. In fact, as the example of official development assistance showed, some studies control for spatial dependence, but the lack of theory and justification for the choice of the weighting matrix indicate that authors are not consciously analysing spatial contagion.

Compared to monadic data, modelling spatial dependence in dyadic data is somewhat more complex. This is not only because there are five different forms of spatial contagion in directed dyads, but also because the weighting matrix allows for a large variety of choices. Both the concrete form of spatial dependence as well as the weighting matrix used need to be theoretically justified, as the results might strongly depend on how spatial dependence modelled. By describing and naming the different forms of spatial dependence both in undirected and directed dyads, Neumayer and Plümper (2010a) have raised the issue in the empirical literature and paved the way for more applications. Even though the calculation of the spatial lags remains computationally intensive, particularly for specific source and specific target contagion, the ado-files for STATA by Neumayer and Plümper (2010b) allow an easy access and provide a convenient way to compute spatial lags for the applied researcher.

References

Baistrocchi, Eduardo (2008), The Use and Interpretation of Tax Treaties in the Emerging World: Theory and Implications, *British Tax Review*, 4, 352-391.

Neumayer, Eric and Thomas Plümper (2010a), Spatial Effects in Dyadic Data, *International Organization***,** 64(1), pp. 145-166.

Neumayer, Eric and Thomas Plümper (2010b), Making Spatial Analysis Operational: Ado-files for Generating Spatial Effect Variables in Monadic and Dyadic Data, *Stata Journal*, 10 (4), pp. 585-605.