Summer 2010

Modeling Foreign Direct Investment as a Longitudinal Social Network

Nathan M. Jensen
Washington University in St Louis, njensen@wustl.edu

Andrew D. Martin
Washington University in St Louis, admartin@wustl.edu

Anton H. Westveld
University of Nevada - Las Vegas, anton.westveld@unlv.edu

Please Do Not Cite Without Permission - Paper in Progress. Presented at the 3rd Political Networks Conference, Duke University, May 20, 2010.

Modeling Foreign Direct Investment as a Longitudinal Social Network

Please Do Not Cite Without Permission — Paper in Progress*

Nathan M. Jensen[a], Andrew D. Martin[b], Anton H. Westveld[c]

[a]Associate Professor of Political Science, Washington University in St. Louis, MO
[b]Professor of Law & Professor of Political Science, Washington University in St. Louis, MO
[c]Assistant Professor of Statistics, Department of Mathematical Sciences,
University of Nevada – Las Vegas, NV†

May 19, 2010

Abstract

An extensive literature in international and comparative political economy has focused on the how the mobility of capital affects the ability of governments to tax and regulate firms. The conventional wisdom holds that governments are in competition with each other to attract foreign direct investment (FDI). Nation-states observe the fiscal and regulatory decisions of competitor governments, and are forced to either respond with policy changes or risk losing foreign direct investment, along with the politically salient jobs that come with these investments. The political economy of FDI suggests a network of investments with complicated dependencies.

We propose an empirical strategy for modeling investment patterns in 24 advanced industrialized countries from 1985-2000. Using bilateral FDI flow and stock data, we examine the nature of the networks in relation to a set of covariates — in particular differences in tax rates between pairs of nations. Our statistical model is based on the methodology developed by Hoff (2005), Westveld (2007), Westveld and Hoff (2009*b*). The model allows the temporal examination of each nation's activity level in investing and attractiveness to investors. Additionally, the model considers the temporal examination of reciprocity between pairs of nations, as well as the notion of clusterability. For both the flow and stock data, there exist a data set based on reports from senders (*out-reported-data*) and a data set based on reports from receivers (*in-reported-data*). We extend the model by treating these two data sets as independent replicates (for the flow and stock data separately), conditional on a mean parameter representing an underlying value of FDI, along with random effects within the variance portion of the distribution of the response that allows for discrepancy between the two data points (in and out data). Using a fully Bayesian approach, we also impute the missing data within a MCMC algorithm used to fit the model.
Keywords: Bayesian inference, foreign direct investment, hierarchical modeling, longitudinal data, social network data.

1 Introduction and Preliminaries

The study of globalization has long been an important part of political science research, yet a recent resurgence scholarship has explored the political implications of the investments of multinational corporations,

*Paper presented at the 3^{rd} Political Networks Conference, Duke University, May 20, 2010. We would like to thank the Center for Empirical Research in the Law at Washington University for supporting our research, as well as Grace S. Chiu and Kevin M. Quinn for discussions during the course of this project.

†Corresponding author: `anton.westveld@unlv.edu`.

or foreign direct investment (FDI). One explanation for this increased interest in international finance by political scientists is that the increase in capital mobility is closely related to politics. Although multinational corporations have long been a part of world politics, there has been a dramatic increase in foreign direct investment in the post-war period and an explosion in FDI flows in the 1990s. Much of this increase in investment activity is caused by economic and technological factors, yet politics plays a central role in facilitating these movements.[1]

Policy choices such as the liberalization of rules on the movement of capital have lead to increased FDI flows.[2] The privatization of state-owned enterprise has lifted numerous transition economies into becoming major recipients of FDI as foreign companies purchase assets and expand production (UNCTAD 2003). Numerous countries have altered tax policies to attract multinational investment[3].

Politics can also deter the entrance of multinational corporations. Governments can provide generous incentives to attract investment, yet once this investment is made, the government can change tax rates, expropriate assets, or engage in other forms of opportunistic behavior that harms firms.[4] Some types of political institutions, such as systems with many veto players can lead to a more stable policy environment for multinationals leading to more FDI. Democratic institutions, beyond increasing the number of veto players, can also provide different costs and benefits for multinationals, affecting patterns of investment[5] and their operations (Jensen 2007, Henisz 2000, 2002).

Not only do politics affect multinationals, but multinationals can also affect politics. At a minimum, investment affects the economy, which can have political ramifications. FDI has numerous direct effects, such as employing workers and paying a wage premium over domestic investors, increasing exports, expanding the tax base (Jensen 2006). FDI can also have spillovers both within and across industries through facilitating technology transfer and other productivity spillovers (Javorcik 2007).

These positive benefits of FDI can change the preference of politicians. Politicians may be wary of increasing tax rates, changing environmental regulations, or making policy decisions that may induce capital flight leading to a “race to the bottom” in government policy. This ability of capital to exit may also reduce a governments “room to maneuver” where government monetary and fiscal policy has limited effectiveness in an open economy(Mosley 2003, Sattler, Freeman and Brandt 2007). Thus, interaction capital can affect the incentives and ability of politicians.[6]

While FDI can have important ramifications for the study of politics, theoretical models of FDI are quite complex. For example, the impact of high tax rates on FDI seems obvious at first brush; higher taxes should be associated with less FDI flows. Yet a complex literature has emerged that explores how multinationals, maximizing worldwide profits, respond to tax rates. Related to this, tax rates also have indirect effects on multinationals, where governments may use corporate tax revenues to finance public goods such as infrastructure that benefits multinationals.

More generally, theoretical models of foreign direct investment are based on market imperfects. Multinational corporations are attempting to maximize global profits, usually requiring firms to produce and sell products and services around the world, yet there are numerous way to achieve these goals. For example, to sell automobiles to Chinese consumers, U.S. producers can simply export autos made in Detroit or contact

[1]See Blonigen (2005) for a review of the economic determinants of FDI flows.

[2]See Quinn and Inclán (1997) and Simmons and Elkins (2004) for a discussion of financial liberalization. Also UNCTAD (2003) for an overview of of the policy determinants of FDI.

[3]Although some scholars have found evidence of tax rates affecting firm operations (Devereux and Griffith 1998, 2003, Hines 1999, Mooij and Ederveen 2003, Mutti 2003) the responsiveness of FDI to tax rates remains an open question (Blonigen 2005). Tax competition across countries—making changes to corporate tax policies and offering generous tax incentives (Li 2006)—remains a controversial area of research in political science Swank and Steimo (2002), Hayes (2003), Basinger and Hallerberg (2004), Jensen (2006), Franzese and Hayes (2006, 2007).

[4]Thee classic work on this relationship is referred to as obsolescing bargaining (Vernon 1971). See Swank and Steimo (2002) and Alfaro, Kalemli-Ozcan and Volosovych (2005) for a discussion of how risk affects FDI flows.

[5]Much work has identified how political institutions affect the risk environment for FDI. See Henisz (2000), Henisz (2002) and Jensen (2003), Jensen (2006) for a discussion of how veto players can reduce political risks for multinationals. See Oneal (1994), Jensen (2003), Jensen (2004) and Li and Resnick (2003) for an exploration of how democratic institutions affects the risk environment for multinationals.

[6]Multinationals can also directly impact politics through campaign contributions, lobbying, or other direct political activities.

a Chinese manufacturer and license the Chinese firm to make the automobiles. Both strategies require no capital investment, thus no FDI. Yet numerous political and economic factors may induce the U.S. producer to invest in a Chinese production facility. Low wage rates and high barriers to trade could make producing autos in the U.S. costly relative to Chinese production. U.S. firms can still license a Chinese manufacture to produce the autos, although this requires a domestic firm with the technical expertise and access to capital. A U.S. firm may choose to invest rather than license because the US firm has access to capital and know-how. These firms could also be induced to invest to take advantage of Chinese tax incentives, or to manage sensitive technology that could be stolen by the Chinese firm.

These examples are not caveats to FDI decisions, the theory of foreign direct investment focuses on FDI as a *response* to market imperfections. According to Dunning's (1981) OLI framework, MNCs are driven to make physical investments abroad for ownership, location, or internalization reasons. Ownership reason for investment are when a foreign firm has ownership advantages over domestic firms, such as having superior access to capital or global brand recognition (Nike, Intel, etc). Simply having advantages over domestic firms doesnt always induce investment, foreign firms could utilize these advantages to produce goods at home and export them abroad. Location factors capture the factors that such as services that require close physical proximity to customers (restraints), goods that entail high transportation costs (bricks), or government policies that make trade prohibitive. Location factors may not be enough to induce investment in all cases, a foreign firm can simply license a domestic firm to produce for a firm in a form of franchises or license agreement. Internalization factors capture the motivations for multinationals to invest abroad as a mechanisms to keep operations internal to the firm, rather than outsourcing or licensing. One of the main motivations for this is that a firm may have an intangible asset that is difficult to transfer to other firms such as a strong corporate culture or the firm is wary of transferring a sensitive technology to another firm.

Thus, important questions, such as how FDI responds to tax rates, are complicated and require careful empirical analysis. Although there are conditions where high tax rates may detract FDI, it also could induce FDI (e.g. if foreign firms are better at avoiding taxes than domestic firms).

Unfortunately, data quality problems plague FDI research. Although most statistical agencies measure FDI as foreign capital coming in the form of equity, reinvested earnings, and intercompany debt, the OECD definition of FDI is more conceptual than technical.

> Foreign direct investment reflects the objective of obtaining a lasting interest by a resident entity in one economy ("direct investor") in an entity resident in an economy other than that of the investor ("direct investment enterprise"). The lasting interest implies the existence of a long-term relationship between the direct investor and the enterprise and a significant degree of influence on the management of the enterprise. Direct investment involves both the initial transaction between the two entities an all subsequent capital transactions between them and among affiliated enterprises, both incorporated and unincorporated (OECD 1996).

The OECD goes on to discuss this over the next 56 pages, clearly illustrating the complexity of defining and measuring FDI. Although the OECD attempts to provide guidance and coordination on the coding of FDI, there are major differences in how countries tabulate FDI. The OECD identifies these major issues (OECD 1996):

1. Timing: Some countries record FDI based on the calendar year and others report based on the accounting year.

2. Reporting: Many OECD countries have mandatory FDI reporting requirements, while some do not.

3. Survey methodology: There are major differences in the surveying methodology across countries.

4. Sampling: Some countries require all firms to report FDI activity, while other countries exempt some small and medium sized enterprises.

In this paper we identify a strategy for exploring patterns of FDI flows and stock. The canonical empirical models in the literature are limited in their capacity to capture network dependencies, and are thus limited in their ability to understand these complicated dynamics. The focus of the project is empirical, focusing on estimating the network effects of FDI inflows and outflows, as well as inward and outward stock positions. By understanding and modeling these networks, this paper can contribute to substantive political economy questions. This paper also addresses the profound data quality problems by imputing missing data within the MCMC algorithm and treating the reported inflow and outflow data, as well as inward and outward as independent replicates of the true value. Finally, the paper makes use of state-of-the-art models to deal with network and temporal dependencies (Hoff 2005, Westveld 2007, Westveld and Hoff 2009*b*,*a*).

2 Data and Exploratory Data Analysis

We investigate the patterns of foreign direct investment (FDI) flows among countries belonging to the Organization for Economic Co-operation and Development (OECD) from 1985-2000, along with a set of covariates which describe a standard gravity model. The FDI data were obtained from the OECD's *International Direct Investment by Country Volume 2009* and consist of four data sets:

1. Flow of FDI: a.) inflows and b.) outflows;

2. Stock of FDI: a.) inward position and b.) outward position.

The in-reported-data sets consist of the reports from every OECD member nation on the amount of FDI their nation received from other nations in a given year (flow), as well as the current total FDI from each nation (stock). The out-reported-data sets are similar, but each nation reports their outflows of FDI to other nations in a given year, and the total FDI currently invested in other nations (stock). This design means that there are two pieces of information for the amount of FDI flow or stock from OECD member i to OECD member j ($i \xrightarrow{t} j$) in a given year t — the out-reported information from i and the in-reported information from j. This leads to modeling data of the form $y^{in}_{i,j,t}$ and $y^{out}_{i,j,t}$. Since each country uses their own criteria, has differing incentives as to what to report, and statistical governmental services, statistical methodologies, etc., the values based upon the out-reported-data and in-reported-data values are generally different and can be considered as two independent opinions (estimates) of the *true* value of FDI (flow or stock). For example in Table 2, the bilateral flows between Denmark and France for 1987 are presented. They are two main points to consider: The first is that FDI flows do not have to be positive. In a given year a country can invest or dis-invest from another country. Secondarily, the outflow and inflow values are not equivalent. In 1987, France estimated its investment in Denmark to be -9.2 (nominal), while Denmark estimated its investment from France to be 6.6.[7] Boxplots of the difference between the real values of in-reported-data and out-reported-data from country i to country j ($y^{in}_{i,j,t} - y^{out}_{i,j,t}$) can be seen in Figure 1[8]. As we might expect, the medians for each boxplot is around zero, suggesting somewhat similar estimates. However as the figure indicates, there does exist some large discrepancies between the in-reported-data and out-reported-data values. For example, In the year 2000 based on the the outflow data, the FDI from the United Kingdom to Germany was 172,210 million US dollars; however, based on the the inflow data, the FDI from the United Kingdom to Germany was only 50,250 leading to a difference of -121,959.

Another interesting fact about the data is the substantial amount of missing values, 42% (flow) and 52% (stock). Since countries self-report their values, many countries do not provide estimates on their in-reported-data and/or out-reported-data of FDI to the OECD. Figure **??** presents 4 different plots to examine the missingness in the data. The first column of the figure depicts the percent missing over time for the flow and stock data. As might be expected, there is a general decline in non-reporting over the period. Additionally, there are more missing values in the stock data compared to the flow data. The second column

[7]The monetary values in this paper are in millions of US dollars. An inflator using the CPI-All Urban Consumers data was calculated to set the amounts into real values based on the year 2000. The CPI data can obtained from the following: *http://www.bls.gov/data/home.htm*. Note: this CPI data is used in the BLS inflation calculator: *http://data.bls.gov/cgi-bin/cpicalc.pl*.

[8]There is a large amount of missing data in these data sets, which will be discussed subsequently. The figure is based upon complete information for the pair $y^{in}_{i,j,t}, y^{out}_{i,j,t}$.

Table 1: The bilateral investment between Denmark and France in 1987. The FDI is presented in millions of $US and the nominal values are the *raw* data obtained from the OECD. The CPI-All Urban Consumers from the US Bureau of Labor Statistics was used to place the nominal amounts into real values based on the year 2000.

From	To	Year	Nom FDI	Real FDI	Type	Reporting Country
Denmark	France	1987	31.4	47.598	Inflow	France
Denmark	France	1987	23.2	35.168	Outflow	Denmark
France	Denmark	1987	-9.2	-13.946	Inflow	Denmark
France	Denmark	1987	6.6	10.005	Outflow	France

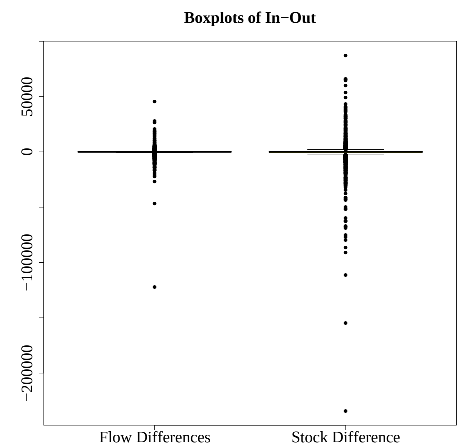


Figure 1: The difference between in-reported-data and out-reported-data values (i.e. $y^{in}_{i,j,t} - y^{out}_{i,j,t}$) for the FDI from country i to country j.

present the percent of missing (non-reported) values for each nation. Based on the flow data, it can be seen that over the period (1985-2000) that Belgium and Luxembourg did not report their FDI values at all. Additionally, Canada and Greece did not report more than 80% of their values. However, we can also see that Germany reports almost all of its values. Next examining the stock data, Belgium, Ireland, and Spain did not report any of their values over the period. Again, Greece did not report more than 80% of it values, however Canada has only about 25% of its values missing. Also, Germany and the US have the lowest percent of missing values.

2.1 Transformations of the Response and Covariates

Next, we consider the distribution of FDI for the OECD countries. Figure 3 shows that the data are heavily right skewed. A possible statistical modeling approach to deal with the skewness is to shift the data to the positive scale and take the log (note that there are far more negative values in the flow data compared to the stock data). Blonigen (2005) states that a shifted log transform is a standard approach for modeling FDI and has appeal since it "is the typical practice with 'gravity' models". However, we suggest a quarter-root transform extended to the whole real, which is a particular version of a broader family of Box-Cox transformations extended to the whole real line (Yeo and Johnson 2000). In particular, consider

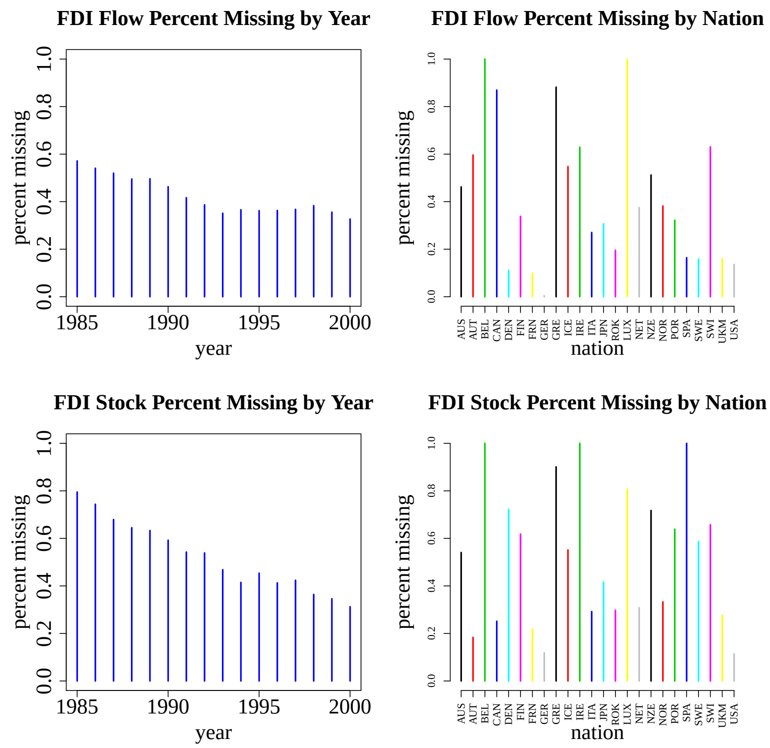


Figure 2: The percent of missing values in the OECD inflows data.

the following two transformations (shifted and pos-neg quarter-root transform)[9]:

$$f(x)^{\text{shifted}} = \log(x - \min(x) + 1) \qquad \min(x) \leq 0,$$

$$f(x)^{\text{pos \& neg 1/4-root}} = \begin{cases} x^{1/4} & x \geq 0 \\ -(-x)^{1/4} & x < 0 \end{cases}.$$

Figure 4 depicts both transforms over the range of -100 to 100. From the figure we notice that the quarter-root transform has a similar same shape as the log transform on the positive scale. On the negative scale the quarter root transform is a mirrors its shape on the positive scale. Next, in Figure 5 we apply the transforms to the FDI flow data. Perhaps surprisingly, the shifted log transform does not remove the skewness. Since the skewness may be due to the shift leading to large positive values, the transform was tried again but first the values of FDI flows were divided by 1000. While this does a better job with the skewness (i.e. the density looks fairly symmetric) we notice that the density has extremely long tails. However, examining the positive-negative quarter-root transform we are actually able to see a more ‘full’ somewhat symmetric density over the range of values. It is important to note that we are not trying to obtain a normal distribution, just to achieve relative symmetry and ‘fullness’ over the range of values. The multiple modes suggest different groupings among the data. It is exactly these groupings that this paper hopes to capture based on the heterogeneity among the network of OECD members[10]

[9]The quarter-root transform has been suggested by Julian Besag for a wide range of applications. A discussion on this can be found at `http://www.stat.columbia.edu/~gelman/blog/`.

[10]As a final note we would like to mention that we found numerous errors in the data. For example, in 2002 Sweden had 0.5

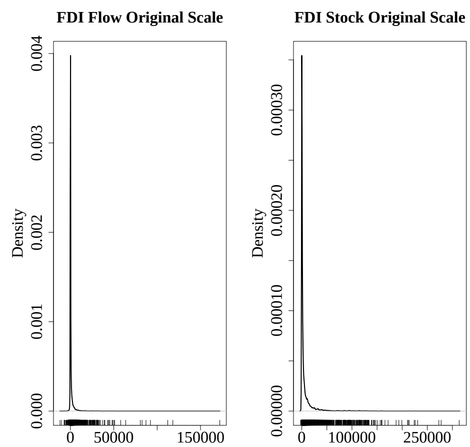


Figure 3: The densities of FDI (flow and stock) for the OECD countries are extremely skewed to the right.

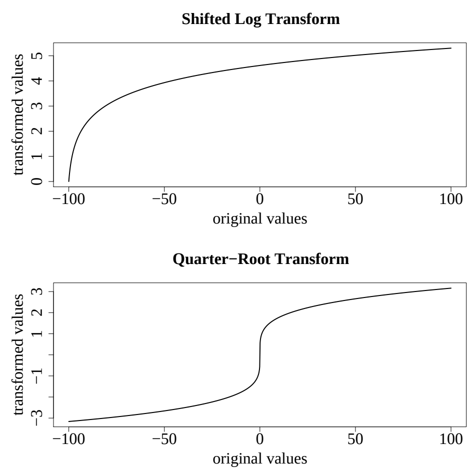


Figure 4: The shifted log and modified log transforms for a sequence from -100 to 100.

In this paper, we examine the relationship between FDI (flow and stock) with a set of covariates that typically are thought of as representing a 'gravity model'. In addition to those covariates, we also consider the tax rate differential between the sending and receiving country, which represents the cost of FDI of the sending country. In particular, the set of covariates and their associated transforms are presented in Table 2.1. Note that for all the covariates in the table, except for the tax rate differential, the log transform was applied (all of those values are positive)[11]. Additionally, the real values for GDP in US dollars were used with 2000 being the base year.

million US dollars worth FDI into itself in the outflows data set. Also we have a concern that some countries miss-interpret NAs and zero FDI.

[11]In future work, this will be compared to quarter-root transformations for the covariates.

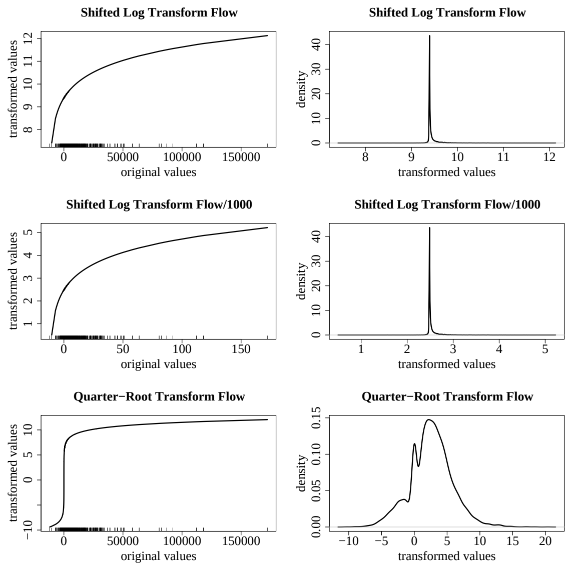


Figure 5: The shifted and modified log transforms applied to the FDI data.

Table 2: Covariates used in the analyses.

Covariate	Transform	Abbreviation
GDP per capita of sender	log(x)	log.gdp.pc.sen
Population of sender	log(x)	log.pop.sen
GDP per capita of receiver	log(x)	log.gdp.pc.rec
Population of receiver	log(x)	log.pop.rec
Distance between the sender and receiver	log(x)	log.dist
Statutory corporate tax rate of the sender minus the receiver	none	tax.sen - tax.rec

2.2 Independent Opinions

As previously mentioned, for each relationship $i \xrightarrow{t} j$, for either the flow or stock data, we have information from j based on the in-reported-data set and information from i based on the out-reported-data set $(y^{in}_{i,j,t}, y^{out}_{i,j,t})$. Since each country uses their own criteria and statistical services to estimate the in-reported and out-reported values for their country, the data $(y^{in}_{i,j,t}, y^{out}_{i,j,t})$, may be thought of as independent replicates since they come from fairly independent sources. One concern however, is that countries may be systematically over or under report their values of FDI. In Figure 6 boxplots are presented for FDI for the reported inflows and outflows, as well as inward and outward positions (i.e. stock). From the figure, there appears to be a slight decrease in the median of reported inflows compared to reported outflows. While there visually appears to be no difference between the medians of the inward and outward positions.

Figure 7 presents scatter plots of the four responses based on a random sample of 2,000 cases. If the in-reported data values are similar to the out-reported data values then we would expect a positive linear relationship between the inflow and outflow values, as well as the inward and outward positions. Examining the flow data, for a large portion of the points there appears to be agreement on the value of FDI flows; however for a substantial portion of the data there appears to be considerable disagreement. As for the stock data, the amount of disagreement is minimal. In order to account for reporting discrepancies, as depicted in

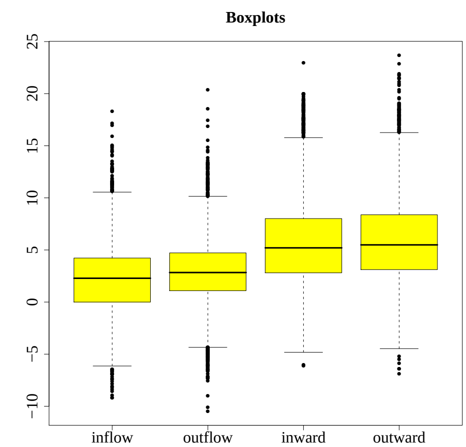


Figure 6: Reported inflows and outflows, as well as inward and outward positions.

Figures 8 and 9, we will consider a model that assumes an underlying level of FDI ($\theta_{i,j,t}$) that incorporates a measure of discrepancy between the reported values ($\sigma^2_{i,jt}$). Additionally, this underlying value of FDI will be assumed to have a temporal network structure.

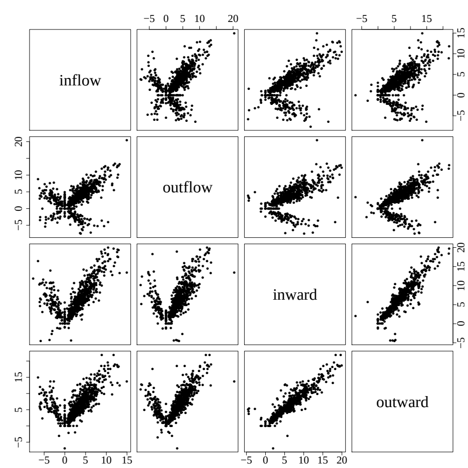


Figure 7: Scatter plots for the four responses based on a random sample of 2,000 data points.

$$
\begin{tikzcd}
y^{in}_{i,j,t-1} \arrow[d, dash] & y^{in}_{i,j,t} \arrow[d, dash] & y^{in}_{i,j,t+1} \arrow[d, dash] \\
\theta_{i,j,t-1} \arrow[r, dash] \arrow[d, dash] & \theta_{i,j,t} \arrow[r, dash] \arrow[d, dash] & \theta_{i,j,t+1} \arrow[d, dash] \\
y^{out}_{i,j,t-1} & y^{out}_{i,j,t} & y^{out}_{i,j,t+1}
\end{tikzcd}
$$

Figure 8: Latent value of FDI is represented by θ.

$$
\begin{tikzcd}
y^{in}_{i,j,t-1} \arrow[d, dash] & y^{in}_{i,j,t} \arrow[d, dash] & y^{in}_{i,j,t+1} \arrow[d, dash] \\
\sigma^2_{i,j,t-1} \arrow[d, dash] & \sigma^2_{i,j,t} \arrow[d, dash] & \sigma^2_{i,j,t+1} \arrow[d, dash] \\
y^{out}_{i,j,t-1} & y^{out}_{i,j,t} & y^{out}_{i,j,t+1}
\end{tikzcd}
$$

Figure 9: Discrepancy of reported FDI values are represented by σ^2.

3 Modeling Strategy and Results

In order to model the FDI data, either the flow or the stock, we will consider the following:

$$
\begin{aligned}
y_{i,j,t,e} &= \theta_{i,j,t} + \epsilon_{i,j,t,e}, \\
\epsilon_{i,j,t,e} &\sim \text{normal}(0, \sigma^2_{i,j,t} = \exp(\kappa_{i,j,t} + \nu_t)); \quad \kappa_{i,j,t} = \kappa_{j,i,t}, \\
\kappa_{i,j,t} &\sim \text{normal}(0, \sigma^2_\kappa), \\
& i, j \in \{1, \ldots, A\}, i \neq j; t \in \{1, \ldots, T\}; e \in \{\text{in-reported}, \text{out-reported}\}.
\end{aligned} \tag{1}
$$

In the model, θ represents a latent level of investment, while σ^2 allows for discrepancy between reported values through a random effects parameter κ. Based upon the work of Gill and Swartz (2001), Westveld (2007), Westveld and Hoff (2009*b*,*a*), we relate the latent level of investment with a set of covariates and account for temporal and pairwise network dependency:

$$
\begin{aligned}
\theta_{i,j,t} &= x'_{i,j,t}\beta + \delta_{i,j,t} \\
\delta_{i,j,t} &= s_{i,t} + r_{j,t} + g_{i,j,t}.
\end{aligned} \tag{2}
$$

In this model, $x'_{i,j,t}\beta_t$ is a fixed effect expressing the mean for $\theta_{i,j,t}$, while the error term $\delta_{i,j,t}$ is decomposed into a set of mean-zero Gaussian random effects. This linear decomposition consists of a sending effect $s_{i,t}$, a receiving effect $r_{j,t}$, and a residual error term $g_{i,j,t}$. For a fixed t, the network dependencies can be characterized by specifying covariance structures for the random effects in (2).[12]

$$
\begin{pmatrix} s_{i,t}, r_{i,t} \end{pmatrix}' = \Phi_{sr} \begin{pmatrix} s_{i,t-1}, r_{i,t-1} \end{pmatrix}' + \epsilon_{sr,t}, \tag{3}
$$

$$
\begin{pmatrix} g_{i,j,t}, g_{j,i,t} \end{pmatrix}' = \Phi_{gg} \begin{pmatrix} g_{i,j,t-1}, g_{j,i,t-1} \end{pmatrix}' + \epsilon_{gg,t}, \quad \text{where}
$$

$$
\Phi_{sr} = \begin{pmatrix} \phi_s & \phi_{sr} \\ \phi_{rs} & \phi_r \end{pmatrix}, \quad \Phi_{gg} = \begin{pmatrix} \phi_g & \phi_{gg} \\ \phi_{gg} & \phi_g \end{pmatrix},
$$

and $\epsilon_{sr,t}$ and $\epsilon_{gg,t}$ are independent mean-zero bivariate normal vectors with covariance matrices Γ_{sr} and Γ_{gg}:

$$
\Gamma_{sr} = \begin{pmatrix} \gamma_s^2 & \gamma_{sr} \\ \gamma_{sr} & \gamma_r^2 \end{pmatrix}, \quad \Gamma_{gg} = \begin{pmatrix} \gamma_g^2 & \lambda_{gg}\gamma_g^2 \\ \lambda_{gg}\gamma_g^2 & \gamma_g^2 \end{pmatrix}. \tag{4}
$$

[12] In the current manuscript a Markov temporal structure is considered, however based on Westveld (2007), Westveld and Hoff (2009*a*) a more general temporal can be considered. Additionally, **?** provide an interesting example of the use of random effects for modeling dependenices among votes within the U.S. Supreme Court.

The resulting covariance matrix for the vector $sr_i = (s_{i,1}, r_{i,1}, \ldots, s_{i,T}, r_{i,T})'$ can be written as

$$Cov(sr_i) = \Sigma_{sr} = \begin{pmatrix} \Sigma_{sr}(0) & \Sigma_{sr}(1) & \cdots & \Sigma_{sr}(T-1) \\ \Sigma_{sr}(1)' & \Sigma_{sr}(0) & \cdots & \Sigma_{sr}(T-2) \\ \vdots & \vdots & & \vdots \\ \Sigma_{sr}(T-1)' & \Sigma_{sr}(T-2)' & \cdots & \Sigma_{sr}(0) \end{pmatrix},$$

where $\Sigma_{sr}(d)$ depends on Φ_{sr}, Γ_{sr} and the time lag d. The covariance matrix of the vector $g_{[i,j]} = (g_{i,j,1}, g_{j,i,1}, \ldots, g_{i,j,T}, g_{j,i,T})'$ has a similar block Toeplitz structure, which we write as $Cov(g_{[i,j]}) = \Sigma_{gg}$, and is made up of the blocks $\{\Sigma_{gg}(0), \ldots, \Sigma_{gg}(T-1)\}$.

As discussed in Westveld (2007), Westveld and Hoff (2009*b*,*a*), the model accounts for network notions of activity through the sending and receiving effects, and reciprocity through the correlation based on $\Sigma_{gg}(0)$. Additional network ideas such as clusterability through a latent space as outlined in Hoff, Raftery and Handcock (2002), Hoff (2005), Ward and Hoff (2007) can be added to the model by expanding $\delta_{i,j,t}$ to include a particular interaction terms:

$$\delta_{i,j,t} \quad = \quad s_{i,t} + r_{j,t} + u'_{i,t} v_{j,t} + g_{i,j,t},$$

where $u_{i,t}$ and $v_{j,t}$ are $k \times 1$ dimensional vectors (for this paper we will take $k = 2$). These random effects are assumed to follow independent mean zero Gaussian distributions with the following covariances:

$$Cov(u_{i,t}) = \begin{pmatrix} \sigma^2_{u1} & 0 \\ 0 & \sigma^2_{u2} \end{pmatrix}, \quad Cov(v_{i,t}) = \begin{pmatrix} \sigma^2_{v1} & 0 \\ 0 & \sigma^2_{v2} \end{pmatrix}.$$

It is important to note that unlike the random effects pairs $s_{i,t}, r_{i,t}$ and $g_{i,jt}, g_{j,i,t}$ the random effects $u_{i,t}, v_{i,t}$ are not correlated over time. However, the $theta$'s and thus the y's are dependent over time based on $s_{i,t}, r_{i,t}$ and $g_{i,jt}, g_{j,i,t}$ random effects. Additionally, as all the random effects in δ have mean zero and a covariance which does not depend on time, they are assumed to be stationary. Again, however, the θ's are not stationary as their mean changes over time, and the y's are not stationary since their mean and variances change over time.

In order to estimate the parameters in this hierarchical it is natural to consider a Bayesian approach. Based on diffuse conjugate or semi-conjugate priors a Markov chain can be constructed for the parameters. In particular the parameters were updated in the following blocks:

$\theta_{i,j,t}, \theta_{j,i,t} \vert \cdot$	via Gibbs sampling
$\beta_1, \ldots, \beta_T \vert \cdot$	via Gibbs sampling
$s_{i,1}, r_{i,1}, \ldots, s_{i,T}, r_{i,T} \vert \cdot$	via Gibbs sampling
$u_{i,1}, v_{i,1}, \ldots, u_{i,T}, v_{i,T} \vert \cdot$	via Gibbs sampling
$\Phi_{sr} \vert \cdot$	via semi-conjugate Gibbs proposal with Metropolis-Hastings correction (Westveld and Hoff 2009*b*)
$\Gamma_{sr} \vert \cdot$	via semi-conjugate Gibbs proposal with Metropolis-Hastings correction (Westveld and Hoff 2009*b*)
$\Phi_{gg} \vert \cdot$	via semi-conjugate Gibbs proposal with Metropolis-Hastings correction (Westveld and Hoff 2009*b*)
$\Gamma_{sr} \vert \cdot$	via semi-conjugate Gibbs proposal with Metropolis-Hastings correction (Westveld and Hoff 2009*b*)
$\sigma^2_{u1}, \sigma^2_{u2}, \sigma^2_{v1}, \sigma^2_{v2} \vert \cdot$	via Gibbs sampling
$\nu_t, \kappa_{i,j,t} \vert \cdot$	via Adaptive Metropolis sampling (Roberts and Rosenthal 2008)
$y_{i,j,t,e}$ (missing values)$\vert \cdot$	via Gibbs sampling

3.1 Flow of FDI

The main focus of this paper is on the flow of FDI. A Markov chain of 29,000 iterations was generated, the first 19,000 of which were dropped to allow for burn-in. Parameter values were saved every 5th scan, resulting in 2,000 samples with which to approximate the joint posterior distribution.[13]

The coefficients for the covariates are presented in Figure 10. The blue bars are 95% credible intervals for the parameters and the black dots are the estimated medians. For our decision rule, heuristically, we will consider intervals which do not contain zero to be a statistically significant covariate. Over the period, the populations of the sender and receiver countries have a positive impact on FDI flows. While the distance between two countries, and for some periods the GDP per capita of the receiver have a negative impact on FDI flows. As for the main question of interest, overall the difference in the tax rate does not appear strongly influence FDI flows (10 of the 16 intervals contain zero). However, if we are more specific, as all the medians, except for one, are greater than zero, overall there is more than 50% chance that the value of the coefficient will be greater than zero. In fact, the probabilities that $\beta_t^{\text{tax.sen-tax.rec}} \geq 0|y$ for $t \in \{1985, \ldots, 2000\}$ are $(0.853, 0.612, 0.954, 0.909, 0.899, 0.961, 0.963, 0.968, 0.819, 0.479, 0.68, 0.588, 0.678, 0.996, 0.995, 0.767)'$.

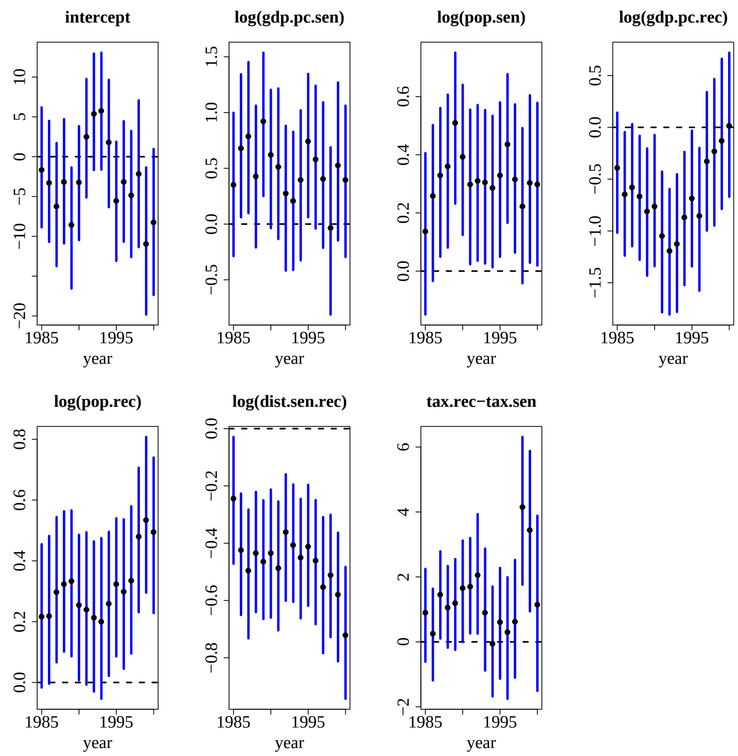


Figure 10: 95% credible intervals

In Figure 11 the means of the bivariate posterior distributions of the sending and receiving effects for each country are plotted. As might be expected, we see in each plot that there exists a strong positive relationship between sending and receiving FDI and that the relative positions of the nations change slightly over the four time points. Note that the United States, Netherlands are located at the top right corner of these plots and thus are the most active nations in the network, after accounting for the covariate information. On the other hand, nations such New Zealand, Greece, and Korea are the least active.

[13]Currently, a longer run of the chain is being conducted.

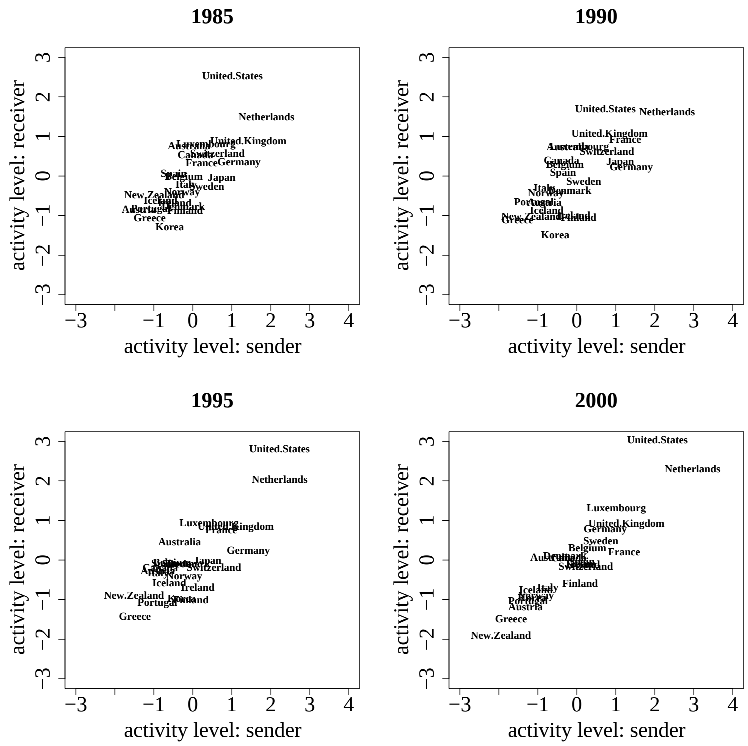


Figure 11: Activity space: sending and receiving

In Figure 12 the means of the bivariate posterior distribution for both the u_t and v_t latent spaces are plotted. Based on the derivation in Chiu and Westveld (2010), a Procrustean transform was applied to the posterior distributions of u_t and v_t for each t based on the posterior means of $u_{t=1}$ and $v_{t=1}$. The idea of the latent space is that actors close together in the space are more similar (similar preferences) than those farther apart. For example, examining the first panel of Figure 12, the United States and United Kingdom are similar in their preference for investing while accounting for covariate information and activity level. In contrast the United States and Ireland have different preferences for investing. Examining the second panel on the first row of the figure, it can be seen that the preference for investing in Norway and Germany are similar but different from Ireland.

Since we assumed a stationary structure, the marginal covariances $\Sigma(0)_{sr}$ and $\Sigma(0)_{gg}$ are estimated to be the same for each year. Based upon the posterior samples for $\Phi_{sr}, \Gamma_{sr}, \Phi_{gg}$, and Γ_{gg} we can compute the posterior distributions of $\Sigma(0)_{sr}$ and $\Sigma(0)_{gg}$. The results can be seen in Table 3, which presents the trace plots of the Markov chain along with the 95% credible intervals and posterior medians. Notice that the medians of the posterior distributions for σ_s^2 and σ_r^2 coincide with the spread of the means of the sender-reciever estimates for the nations (Figure 11). Also shown in the table are (1) the median posterior correlation (instead of covariance) between the sending and receiving effects which is 0.785, and (2) the median posterior correlation within a pair of nations is 0.455. The latter suggests some reciprocity among actors in the network within a given year.

We also examine the auto-regressive coefficients to see what affect the previous year has on sending and receiving FDI flows, along with reciprocity for the current year. From Table 4, the medians of the posterior distributions of ϕ_s and ϕ_{sr} are 0.934 and 0.035, respectively. This suggests that the level of sending FDI this year is highly dependent on level of investment sent in the previous but perhaps not dependent on

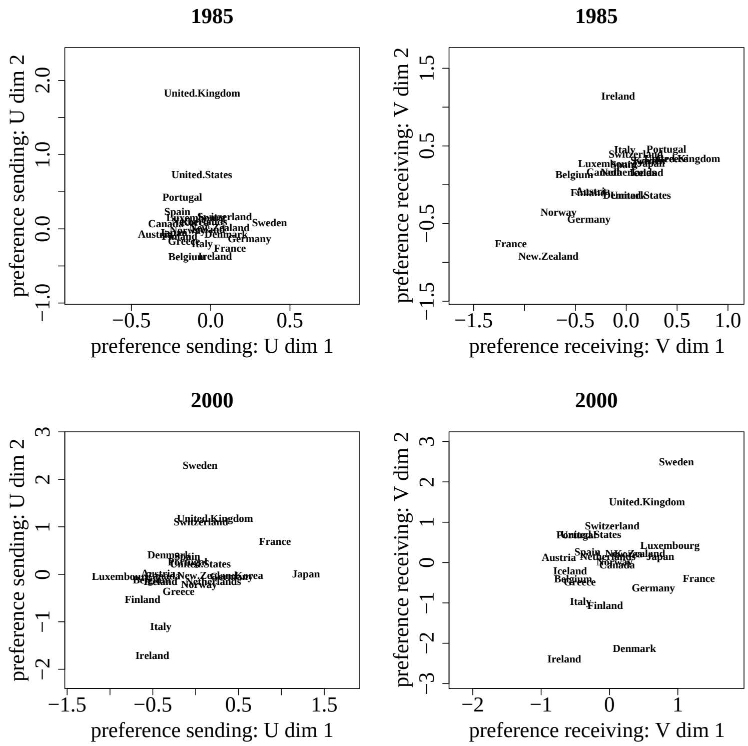
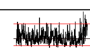
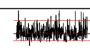
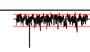
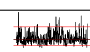
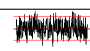
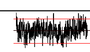
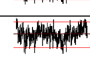


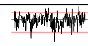
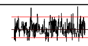
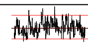
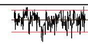
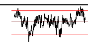
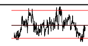
Figure 12: Preference social space: sending and receiving

Table 3: $\Sigma(0)_{sr}$ and $\Sigma(0)_{gg}$

Parameter	Markov chain	2.5%	Median	97.5%
γ_s^2		0.736	1.230	2.150
γ_{sor}		0.508	0.956	1.791
ρ_{sr}		0.628	0.785	0.895
γ_r^2		0.708	1.246	2.273
γ_g^2		1.285	1.475	1.683
γ_{gg}		0.457	0.676	0.872
ρ_{gg}		0.334	0.455	0.562

investment received from the previous year. Comparatively, the medians of the posterior distribution for ϕ_r and ϕ_{rs} are 0.916 and 0.065, respectively. That is, the level of receiving direct investment this year is fairly dependent on the amount of investment received in the previous not that dependent on investment sent in the previous year. As the median of the posterior distribution of ϕ_{gg} is 0.117, we see that a relatively small amount of positive reciprocity in a given year can be explained by the level of reciprocity in the previous year.

Table 4: Φ_{sr} and Φ_{gg} parameter estimates

Parameter	Markov chain	2.5%	Median	97.5%
ϕ_s		0.860	0.934	0.974
ϕ_{sr}		-0.009	0.035	0.101
ϕ_{rs}		0.006	0.065	0.133
ϕ_r		0.850	0.916	0.967
ϕ_g		0.785	0.832	0.862
ϕ_{gg}		0.076	0.117	0.166

Next we would like to examine the discrepancy parameters (κ's). For each nation i at each time point t the mean of the mean of the posteriors $\bar{\kappa}_{i,1,t}, \bar{\kappa}_{i,2,t}, \ldots \bar{\kappa}_{i,A,t}$ was calculated (recall $\kappa_{i,j,t} = \kappa_{j,i,t}$). The results are in Figure 13. Based on the figure it can be seen that Iceland is one of the lowest (reduction in the variance) over several years. In fact, Iceland is below zero on 11 out 16 occasions. On the other hand, France, Germany, Japan are above zero on 11 out 16 occasions. However, it may be conjectured that the percent of missing data through data imputation may be driving the discrepancy pattern. The second panel of the figure examines this question through a scatter plot with a lowess smoother. From the figure there does not appear to be a general relationship between discrepancy and percent missing.

Figure 14 examines some diagnostics based on a random sample of 2,000 points. The first panel of the figure plots the means of the posteriors of the θ's against the means of the posteriors of the fitted values $(x'_{i,j,t}\beta_t + s_{i,t} + r_{j,t} + u'_{i,t}v_{j,t})$. If we believe θ to represent a latent flow, then the model appears to be doing fairly well, except for a slight curvature. The second panel on the first row, examines the means of the posteriors of the residuals against means of the posteriors of the fitted values. There appear to be some outliers and a very slight curvature. The next panel is a Gaussian QQ-plot of the residuals. Through the middle the plot appears to have a slight curvature along with a few potential outliers. Next, Figure 15 exams the the means of the posteriors of the θ's against the observed y's - the inflows data are represented by red dots and the outflow data are represented by blue dots. When the value of y is greater than zero, the model is doing quite well in capturing the observed data. However, when y is less than zero, the model does not do as well. The second panel in the figure examines the imputation of the y's. In order to generate the plot, 100 values from the inflow (red) data and 100 values from the outflow (blue) data were removed before estimation and these were imputed at each scan of the MCMC. The means of the posterior values were used for prediction and are plotted against the observed values. The mean-sqared-error was calculated to be 8.48 on the quarter-root scale, while was 396,173,3 on the original scale. Fitting the same model using the shifted log transform lead to mean-squred-error on the original scale of 812,98,76 — a 200% increase. As noted in a previous iteration of this work, a mixture model is able to correct for the lack of fit when $y < 0$ if the data are observed. However, if the data are missing, this becomes a more difficult problem and is currently being fully investigated.

3.2 Stock of FDI

A full presentation of the analysis of the stock data will not be currently be presented, however Figure 16 presents the 95% credible intervals of the coefficients. As the main scientific question is related to taxes, changes in stock related to tax differentials, except for potentially for the last two years, do not appear to have a strong effect on level of FDI stock. Based upon the diagnostics presented in Figure 17, including the out-of-sample prediction, these results seem quite robust.

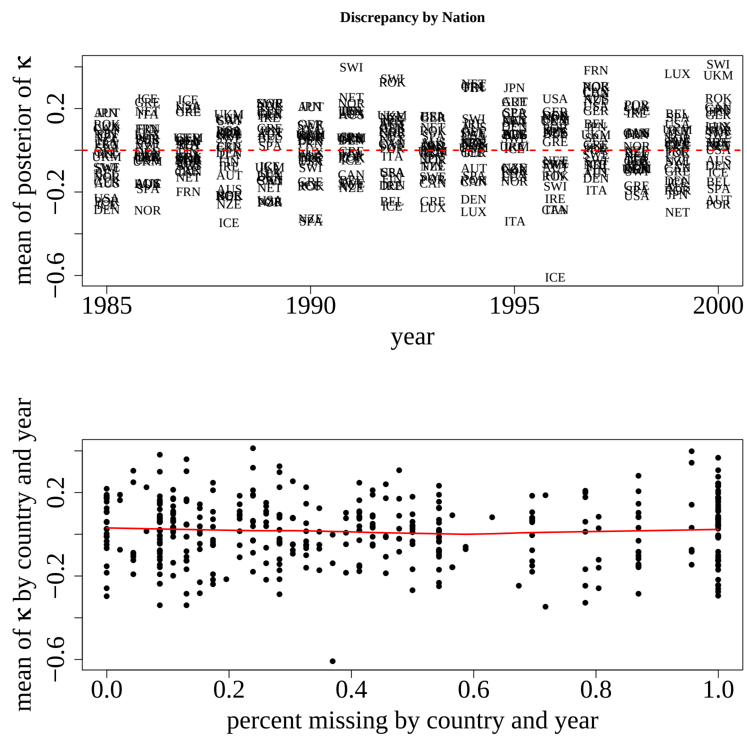


Figure 13: Discrepancy by county and year

4 Discussion

Modeling the network of foreign direct investment is quite complicated, not only due to data limitations, but also due to expected complicated dependencies due to the dynamic network. Existing cross-sectional time series models do not capture network dependencies, and are thus limited in their applicability. Here we provide a dynamic network model that utilizes the work of (Westveld 2007, Westveld and Hoff 2009*b*,*a*). Additionally, we consider a model that accounts for reporting differences in the data. Ultimately we find that the model does a fairly good (but not outstanding) job with the flow data, but still leads to a 200% reduction in the mean-squared-error compared to the shifted log transform in out-of-sample prediction. Additionally, the model does quite well at capturing the stock data.

These empirical results lead to two types of substantive applications to the literature on FDI. First, our empirical work is one of the first studies to estimate the network effects of foreign direct investment. We provide estimates of the sender-receiver effects on both the activity space as well as the preference space for all countries for all years. Thus we have the ability to identify which countries have similar investment profiles and explore how these profiles change over time.

Second, our models allow us to take a nuanced view on the relationship between tax rates and foreign direct investment flows and stock. Specifically, we find that corporate tax rates appear to slightly positively affect FDI flows over the period of study (as the medians are above zero for all years except for one). Additionally, the 95% credible intervals for some years do not contain zero. However, this impact on the flows, does not appear to translate to the stock of FDI, except perhaps for the years 1999 and 2000. However, more investigation into this topic is necessary through modeling improvements. A key to these improvements may be three fold: 1.) jointly model the flow and stock variables, as they are intrinsically linked; 2.) think

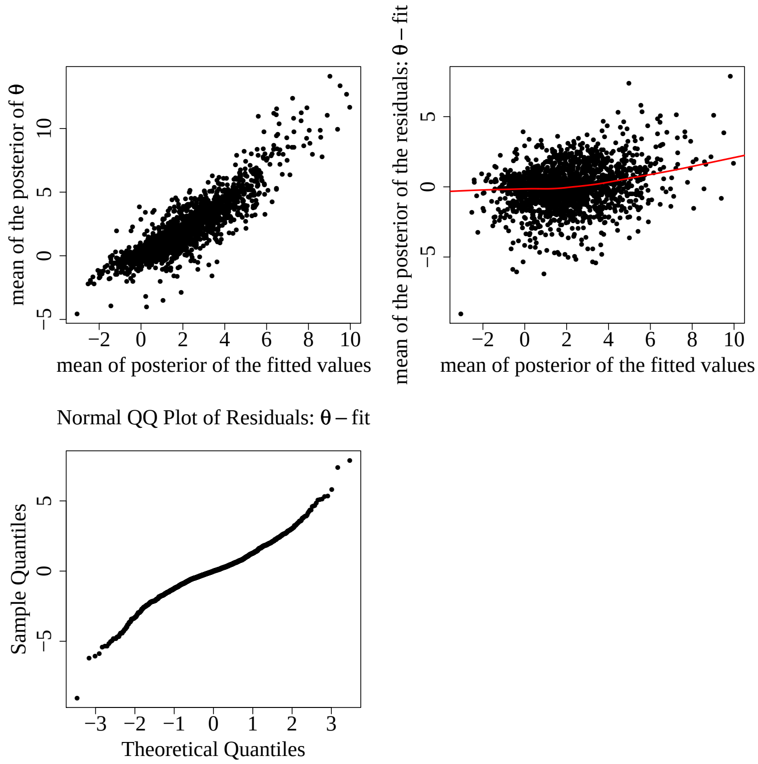


Figure 14: Diagnostics

about other modeling on the level of the response for the flow data; and 3.) consideration of a model that allows for probability statements about whether a particular coefficient is exactly zero or not.

References

Alfaro, Laura, Sebnem Kalemli-Ozcan and Vadym Volosovych. 2005. Why Doesn't Capital Flow from Rich to Poor Countries? An Empirical Investigatio. Technical report NBER Working Paper.

Basinger, Scott J. and Mark Hallerberg. 2004. "Remodeling the Competition for Capital: How Domestic Politics Erases the Race to the Bottom Remodeling the Competition for Capital: How Domestic Politics Erases the Race to the Bottom." *American Political Science Review* 98(2):261–276.

Blonigen, Bruce A. 2005. A Review of the Empirical Literature on FDI Determinents. Technical report University of Oregon and NBER.

Chiu, Grace S. and Anton H. Westveld. 2010. "Modelling Feeding Patterns in Relation to Phylogeny in Food Webs." Working paper-in progress.

Devereux, Michael P. and R. Griffith. 1998. "Taxes and the Location of Production: Evidence from a Panel of U.S. Multinationals." *Journal of Public Economics* 68(3):335–367.

Devereux, Michael P. and Rachel Griffith. 2003. "Evaluating Tax Policy for Location Decisions." *International Tax and Public Finance* 70:107–126.

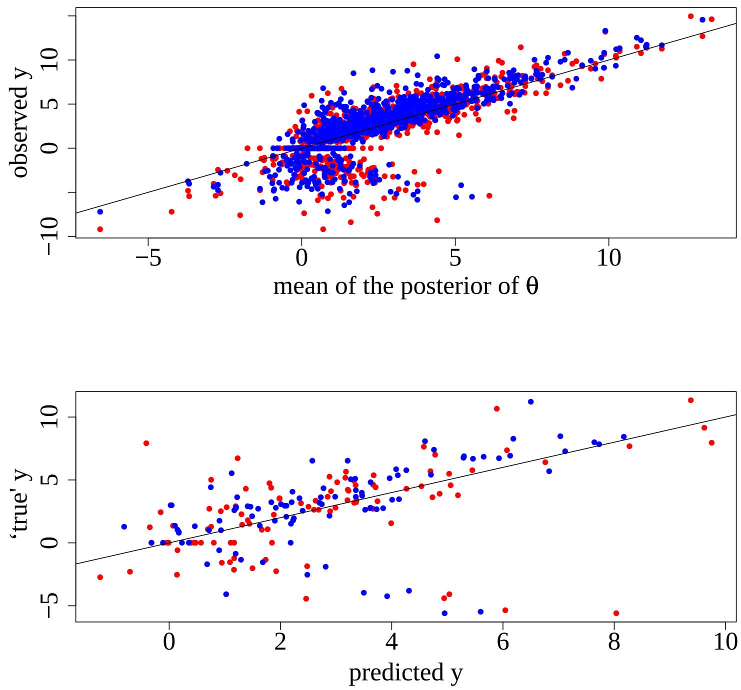


Figure 15: Diagnostics: inflow data (red) and outflow data (blue)

Franzese, Robert J. and Jude C. Hayes. 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):167–89.

Franzese, Robert J. and Jude C. Hayes. 2007. "Spatial-Econometric Models of Cross-Sectional Interdependence in Political-Science Panel and Time-Series-Cross-Section Data." *Political Analysis* 15(2):140–164.

Gill, Paramjit S. and Tim B. Swartz. 2001. "Statistical analyses for round robin interaction data." *The Canadian Journal of Statistics* 29.

Hayes, Jude C. 2003. "Globalization and Capital Taxation in Consensus and Majoritarian Democracies." *World Politics* 56:79–113.

Henisz, Witold J. 2000. "The Institutional Environment for Economic Growth." *Economics and Politics* 12(1):1–31.

Henisz, Witold J. 2002. *Politics and International Investment.* Cheltenham, UK: Edward Elgar.

Hines, James R. 1999. "Lessons from the Behavior Responses to International Taxation." *National Tax Journal* 52:305–322.

Hoff, Peter D. 2005. "Bilinear Mixed-Effects Models for Dyadic Data." *Journal of the American Statistical Association* 100:286 − 295.

Hoff, Peter D., Adrian E. Raftery and Mark S. Handcock. 2002. "Latent Space Approaches to Social Network Analysis." *Journal of the American Statistical Association* 97:1090–1098.

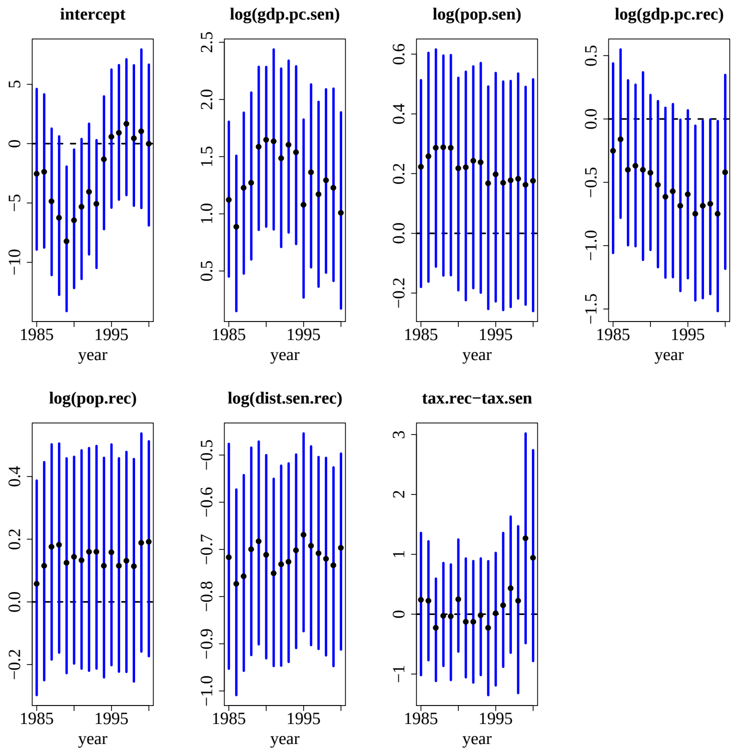


Figure 16: 95% credible intervals

Javorcik, Beata Smarzynska. 2007. Foreign Direct Investment and International Technology Transfer: What Do We Know? What Would We Like to Know? Technical report World Bank Working Paper.

Jensen, Nathan M. 2003. "Democratic Governance and Multinational Corporations: Political Regimes and Inflows of Foreign Direct Investment." *International Organization* 57(3):587–616.

Jensen, Nathan M. 2004. "Crisis, Conditions, and Capital: The Effects of International Monetary Fund Agreements on Foreign Direct Investment Inflows." *Journal of Conflict Resolution* 48(2):194–210.

Jensen, Nathan M. 2006. *Nation-States and the Multinational Corporation*. Princeton: Princeton University Press.

Jensen, Nathan M. 2007. Political Regimes and Political Risk: Democratic Institutions and Expropriation Risk for Multinational Investors. Technical report Washington University in St. Louis.

Li, Quan. 2006. "Democracy, Autocracy, and Tax Incentives to Foreign Direct Investors: A Cross-National Analysis." *The Journal of Politics* 68(1):62–74.

Li, Quan and Adam Resnick. 2003. "Reversal of Fortunes: Democratic Institutions and Foreign Direct Investment Inflows to Developing Countries." *International Organization* 57(1):175–212.

Mooij, Ruud A. Do and Sjef Ederveen. 2003. "Taxation and Foreign Direct Investment: A Synthesis of Empirical Research." *International Tax and Public Finance* 48(2):673–693.

Mosley, Layna. 2003. *Global Capital and National Governments*. Cambridge University Press.

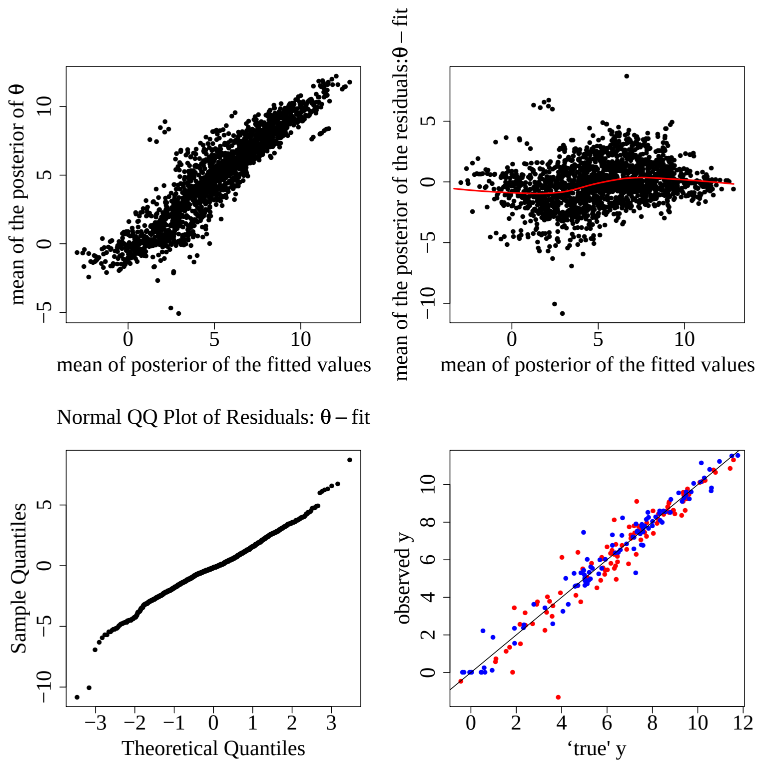


Figure 17: Diagnostics

Mutti, John H. 2003. *Foreign Direct Investment and Tax Competition.* Washington D.C.: Institute for International Economics.

OECD. 1996. *OECD Benchmark of Foreign Direct Investment.* OECD: Paris.

Oneal, John R. 1994. "The Affinity of Foreign Investors for Authoritarian Regimes." *Political Research Quarterly* 47(3):565–588.

Quinn, Dennis P. and Carla Inclán. 1997. "The Origins of Financial Openness: A Study of Current and Capital Account Liberalization." *American Journal of Political Science* 41(3):771–813.

Roberts, Gareth O. and Jeffrey S. Rosenthal. 2008. Examples of Adaptive MCMC. Technical report Department of Statistics, University of Toronto.

Sattler, Thomas, John R. Freeman and Patrick T. Brandt. 2007. "Political Accountability and the Room to Maneuver: A Search for a Causal Chain." *Forthcoming in Comparative Political Studies* .

Simmons, Beth A. and Zachary Elkins. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy." *American Political Science Review* 98(1):171–89.

Swank, Duane and Sven Steimo. 2002. "The New Political Economy of Taxation in Advanced Capitalist Democracies." *American Journal of Political Science* 46(3):642–55.

UNCTAD. 2003. *World Investment Report 2003.* New York: UNCTAD.

Vernon, Raymond. 1971. *Sovereignty at Bay.* New York: Basic Books.

Ward, Michael D. and Peter D. Hoff. 2007. "Persistent Patterns of International Commerce." *Journal of Peace Research* 44(2):157–175.

Westveld, Anton H. 2007. Statistical Methodology for Longitudinal Social Network Data PhD thesis Department of Statistics, University of Washington, Seattle, WA.

Westveld, Anton H. and Peter D. Hoff. 2009*a*. "An Approach for Modeling Longitudinal Social Network Data with General Stationary Covariance Matrices." Working paper-in progress.

Westveld, Anton H. and Peter D. Hoff. 2009*b*. "A Bayesian Mixed Effects Model for Longitudinal Social Network Data." Revision submitted.

Yeo, In-Kwon and Richard A. Johnson. 2000. "A New Family of Power Transformations to Improve Normality or Symmetry." *Biometrika* 87(4):954–959.