# Chapter 5
# Trade-FDI Linkages in a Simultaneous Equations System of Gravity Models for German Regional Data

## 5.1 Introduction

We use a system of simultaneous gravity equations to model German (regional) trade and FDI patterns within the EU27 and to explore correlations among these variables. Whereas predictions from standard trade models of the Heckscher–Ohlin type typically handle both variables as substitutes, recent theoretical contributions in the field of New Trade Theory (NTT) show a more diverse picture when accounting for the growing complexity of investment strategies by multinational enterprises (MNEs), which may follow either horizontal (market-seeking) and/or vertical (cost oriented) investment motives. Depending on the mixture of these two modes, both substitutive and complementary linkages could potentially arise, crucially depending on the chosen model assumptions.[1] Adding on the theoretical literature in solving the trade-FDI puzzle, there is also a steadily increasing stock of empirical contributions, which aim to gain insights to the trade-FDI relationships for individual countries or country groups. Though there is a general tendency for complementary linkages, the empirical literature also gives merely heterogeneous answers to this question. According to Aizenman and Noy (2006), an important aspect to account for in empirical work is to closely interpret the estimation result in light of the chosen country, industry sample and time period.

The research effort spent on solving the trade-FDI puzzle reflects the interest on this subject in the policy debate. As Pantulu and Poon (2003) point out, trade sub-

[1] Markusen (1995), Jungmittag (1995), Zarotiadis and Mylonidis (2005), Helpman (2006) and Blanchard et al. (2008) among others provide detailed surveys of recent theoretical contributions.

This article has been previously published as "Trade-FDI Linkages in a Simultaneous Equations System of Gravity Models for German Regional Data", in: *International Economics*, Vol. 122, pp. 121–162. We kindly acknowledge the permission of the editors to reprint the article in this monograph.

Jointly with Björn Alecke and Gerhard Untiedt. Björn Alecke, Gesellschaft für Finanz- und Regionalanalysen (GEFRA), e-mail: Alecke@gefra-muenster.de; Gerhard Untiedt, GEFRA & Technical University Clausthal, e-mail: Untiedt@gefra-muenster.de.

T. Mitze, *Empirical Modelling in Regional Science*, Lecture Notes in Economics and Mathematical Systems 657, DOI 10.1007/978-3-642-22901-5_5, 

stitutability and replacement effects are sensitive issues in the globalization debate of industrialized countries, linking outward FDI typically to deindustrialization and displacement effects of employment, especially in export-based industries. Thus, for relatively open economies like Germany this analysis may be seen as a very sensitive and important issue. Only few empirical studies have dealt with the German trade-FDI interrelations so far, where the results generally show a substitutive relationship between exports and outward FDI at the national level (see Jungmittag 1995, for selected European countries and the USA between 1973–89 as well as Egger and Pfaffermayr 2004, for a world sample between 1989–99). Accounting for the different historical patterns of unified Germany, an in-depth study of macro-regional differences between East and West Germany may also add a useful new dimension to the trade-FDI debate. This may answer the question in how far political and economic path dependencies in building up trade relations and foreign direct investment stocks may influence the actual internationalization strategies of firms.

To shed more light on the national and regional trade-FDI puzzle, we thus analyze the intra-EU27 trade and FDI patterns for the 16 German federal states (NUTS1-level) based on a panel data set of bilateral region-to-nation trade volumes and FDI stocks between 1993 and 2005.[2] We apply gravity type models in order to identify the driving forces of trade and FDI activity as proposed by the NTT and to gain insight into the likely nature of their interrelation. Econometrically, we estimate both instrumental variable (IV) and non-IV simultaneous equation models accounting for a likely correlation among the individual behavioral equations for trade and FDI. This strategy allows us to identify the underlying nature of the trade-FDI-nexus for Germany by isolating the pairwise effects of trade and FDI on the respective other variable, when controlling for a set of common external factors. Moreover, given the emphasis on the regional modelling perspective, we also put a special focus on a sensitivity analysis of the results with respect to the West and East German macro regions for different EU sub-aggregates.

The remainder of the chapter is organized as follows: Sect. 5.2 sketches the theoretical underpinnings of gravity type model of trade and FDI and also discusses its empirical operationalization. Section 5.3 gives a short literature review with respect to recent theoretical and empirical contributions to analyze trade-FDI-linkages in an international context. Section 5.4 then presents the database and some stylized facts for German trade and FDI patterns within the EU27. Section 5.5 then discusses the time series properties of the variables, the choice of the econometric estimator and our empirical results. Further, robustness checks are performed. Based on our empirical identification strategy, Sect. 5.6 reports the results for the trade-FDI linkages of the German aggregate and regional data. Section 5.7 concludes the chapter.

[2] It would be desirable to have region-to-region trade/FDI data between Germany and the EU27 economies. Unfortunately no such records are available.

## 5.2 Gravity Models of Trade and FDI

### *5.2.1 Theoretical Foundations*

Given its empirical flexibility to model factor flows between regional and national entities in space, the gravity model has a long tradition in the field of international economics (see e.g. Matyas 1997; Feenstra 2004, for a recent overview). The empirical success of the model may be best explained by two facts: It is easy to apply empirically and its results are remarkably good. Starting as a rather ad-hoc empirical specification in the pioneering work of Tinbergen (1962) and Pöyhönen (1963), different scholars have also shown that the model can be derived consistently from theoretical trade models. Whereas earlier work particularly focused on export and import relationships, recent approaches have also adapted the framework to model FDI flow/stock movements motivated by common time features of trade and FDI (see e.g. Brenton et al. 1999). This section is intended to give a short sketch of the model's theoretical foundation and empirical operationalization.

In its fairly simple specification the standard gravity approach models trade between two countries as proportional to the (economic) mass of the countries (typically measured by GDP and population) and inversely related to the distance between them, adopting Newton's law for gravitational forces $GF$ as

$$GF_{ij} = \frac{M_i M_j}{D_{ij}} \quad \text{for } i \neq j, \tag{5.1}$$

where $M_{i(j)}$ are the masses of two objects $i$ and $j$, and $D_{ij}$ is the distance between them. While the first variables proxy supply and demand conditions at home and abroad, the latter serves to measure obstacles to trade. The basic model can be augmented by several other variables and Lamotte (2002) argues that the choice of variables constitutes an important and delicate point, which has to be guided by theoretical and statistical concerns. Looking at its theoretical foundations, the gravity model can arise from a potentially large class of underlying economic structures. Anderson (1979), Helpman (1987) and Bergstrand (1985, 1989) were among the first to show that the gravity model can indeed be derived from a theoretical model. In the trade literature gravity type models based on classical Ricardian models, Heckscher–Ohlin models (see (Deardorff 1998)) and increasing returns to scale models of the NTT have been presented since then. As Henderson and Millimet (2008) summarize, though being different in structure, the models typically have the following common elements: 1) trade separability, which arises when local production and consumption decisions are separable from bilateral trade decisions among locations, 2) the aggregator of differentiated products is identical across locations and is of the constant elasticity of substitution form and 3) trade costs are invariant to trade volumes.

Based on these assumptions and considering a one-sector economy, where consumers have a common elasticity of substitution $\sigma$ among all goods as well as symmetric transportation costs among trading partners, Anderson and van Wincoop (2003) derive a theory consistent gravity model equation as

$$Y_{ij} = \frac{X_i X_j}{X_w}\left(\frac{T_{ij}}{P_i P_j}\right)^{1-\sigma} \quad \text{or:} \quad Y_{ij} = k X_i X_j T_{ij}^{1-\sigma} P_i^{\sigma-1} P_j^{\sigma-1}, \tag{5.2}$$

where $k = 1/X_w$. $Y_{ij}$ is the nominal value of exports from country $i$ to $j$, $X_{i(j)}$ denotes total income for $i(j)$, $X_w$ is world income, $(T_{ij} - 1)$ reflect 'iceberg' transportation (trade) costs and $P_{i(j)}$ are further (multilateral) resistance variables as described by Anderson and van Wincoop (2003).[3] Iceberg transportation costs indicate that $T_{ij}$ units of the product must be shipped to country $j$ in order for one unit to arrive. Feenstra (2004) proposes to model trade costs $T_{ij}$ as a function of distance $d_{ij}$ and other border effects associated with selling from country $i$ to $j$. A similar specification can be used for modelling FDI.

### 5.2.2 *Empirical Operationalization*

We use the gravity model to specify a system of gravity equations for trade and FDI. Here, we have to decide whether to pool the data or use a cross-sectional specification and whether to estimate the gravity model from (5.2) in a log-linearized form. For a detailed discussion of the former point see e.g. Egger (2000), who points out several advantages of the panel data approach.[4] A discussion of the proper functional form in terms of a (log-)linear or non-linear specification is given in Coe and Tamirisa (2002), Henderson and Millimet (2008), as well as Santos Silva and Tenreyro (2006). The latter authors point to the fact that results may be misleading in the presence of heteroscedastic error terms. Since we are dealing with regional data, a correlation of cross-sections may indeed be a potential source of heteroscedasticity. To account for this, we follow Sarafidis and Robertson (2009) and include a set of time dummies, which should at least capture the homogeneous impact of cross-sections to unobserved common factors as one source of heteroscedastic errors. Additionally, Henderson and Millimet (2008) give strong evidence that concerns in the gravity literature over functional form appear unwarranted and that log-linear specifications offer reliable model predictions.[5]

Given the advantages of a panel specification over the cross-section approach, we operationalize the gravity model from (5.2) in line with Cheng and Wall (2002),

[3]In a multi-country framework $X_w$ is defined as $X_w = \sum_{i=1}^{C} X_i$ with $i, j = 1, \ldots, C$ countries.

[4]First, a panel specification catches unobserved heterogeneity in the data caused by time-invariant individual effects (cross-section specific). Second, it allows capturing the relationships between the relevant variables over a longer period and hence is able to identify the role of the overall business cycle phenomenon. Moreover, given the unobserved nature of $P_i$ and $P_j$ in (5.2) a panel data model proxying these effects (for region $i$ and $j$ and/or an interaction term of the form $i \times j$) may thus be a promising alternative to an modelling strategy that tries to directly calculate these resistance variables (see Feenstra 2004, for an overview of different modelling strategies).

[5]The argument raised in Coe and Tamirisa (2002) relates to the problem of missing data due to log-linearization. We take up this point when discussing the data in Sect. 5.4.

Serlenga and Shin (2007) or Egger and Pfaffermayr (2004) in a log-linear way as:[6]

$$y_{ijt} = \alpha + \beta' \mathbf{x}_{ijt} + \gamma' \mathbf{z}_{ij} + u_{ijt} \quad \text{with } u_{ijt} = \mu_{ij} + \nu_{ijt}. \tag{5.3}$$

Here, $y_{ijt}$ represents country $i$'s internationalization activity with respect to country $j$ for time period $t$ (either trade or FDI), with $i = 1, 2, \ldots, N$; $j = 1, 2, \ldots, M$ and $t = 1, 2, \ldots, T$.[7] With regard to the explanatory regressors, $\mathbf{x}_{ijt}$ is a variable vector with variations in three dimensions (home country, host country and time $[x_{ijt}]$), with variation only in time and home country $[x_{it}]$ or time and foreign country $[x_{jt}]$ respectively. Analogously, $\mathbf{z}_{ij}$ is a variable vector of time fixed regressors. $\beta$ and $\gamma$ are vectors of regression coefficients, $\alpha$ is the overall constant term and $u_{ijt}$ is the composite error term including the unobservable individual effects $\mu_{ij}$ (country pair or individual country/region effects) and a remainder error term $\nu_{ijt}$. Typically, the latter two are assumed to be i.i.d. residuals with zero mean and constant variance.

We use a broad set of exogenous control variables in both $\mathbf{x}_{ijt}$ and $\mathbf{z}_{ij}$ to account for any simultaneity bias which arise because of a spurious correlation between trade and FDI when there are common exogenous factors that are affecting both these variables. This allows us to properly isolating the effect of trade and FDI measures on the respective other variables. A common way to run such a identification strategy is to specify the trade and FDI equations and then use the estimation residuals to run a regression as $\lambda_{ijt} = f(\phi_{ijt})$, where $\lambda_{ijt}$ is the residual of the FDI regression (with $ij$ denoting bilateral interaction between country $i$ and $j$, $t$ is the time index) and $\phi_{ijt}$ is the residual of the trade regression (or vice versa). Any significant positive or negative variable coefficient can then be interpreted in favor of non-zero trade-FDI linkages.[8]

Thus, using a log-linear form and variable selection based on both theoretical and statistical concerns, our resulting estimation system can be summarized as follows

$$\begin{aligned}
\log(EX_{ijt}) = {} & \alpha_0 + \alpha_1 + \alpha_2 \log(GDP_{jt}) + \alpha_3 \log(POP_{it}) \\
& + \alpha_4 \log(POP_{jt}) + \alpha_5 \log(PROD_{it}) + \alpha_6 \log(DIST_{ij}) \\
& + \alpha_7 SIM + \alpha_8 RLF + \alpha_9 EMU \\
& + \alpha_{10} EAST + \alpha_{11} BORDER + \alpha_{12} CEEC + \sum_{r=1993}^{2005} \alpha_r t_r,
\end{aligned} \tag{5.4}$$

$$\begin{aligned}
\log(FDIout_{ijt}) = {} & \beta_0 + \beta_1 \log(GDP_{it}) + \beta_2 \log(GPD_{jt}) + \beta_3 \log(POP_{it}) \\
& + \beta_4 \log(POP_{jt}) + \beta_5 \log(PROD_{it}) + \beta_6 \log(DIST_{ij})
\end{aligned}$$

[6] In running the empirical regressions, we also tested for alternative specification and evaluated them in terms of variable significance and post estimation model testing.

[7] Throughout the analysis, $i$ identifies German states, while $j$ represents the EU27 trading partner countries.

[8] Among the earlier contributions to this two-step approach determining trade-FDI linkages are Graham (1999) and Graham and Liu (1998), as well Brenton et al. (1999).

$$+ \beta_7 \log(WAGE_{jt}) + \beta_8 \log(FDIopen_{jt}) + \beta_9 \log(K_{jt})$$
$$+ \beta_{10} SIM + \beta_{11} RLF + \beta_{12} EMU$$
$$+ \beta_{13} EAST + \beta_{14} BORDER + \beta_{15} CEEC + \sum_{r=1993}^{2005} \beta_r t_r, \quad (5.5)$$

$$\log(IM_{ijt}) = \gamma_0 + \gamma_1 \log(GDP_{it}) + \gamma_2 \log(GDP_{jt}) + \gamma_3 \log(POP_{it})$$
$$+ \gamma_4 \log(POP_{jt}) + \gamma_5 \log(PROD_{jt}) + \gamma_6 \log(DIST_{ij})$$
$$+ \gamma_7 SIM + \gamma_8 RLF + \gamma_9 EMU$$
$$+ \gamma_{10} EAST + \gamma_{11} BORDER + \gamma_{12} CEEC + \sum_{r=1993}^{2005} \gamma_r t_r, \quad (5.6)$$

$$\log(FDIin_{ijt}) = \delta_0 + \delta_1 \log(GDP_{it}) + \delta_2 \log(GDP_{jt}) + \delta_3 \log(POP_{it})$$
$$+ \delta_4 \log(POP_{jt}) + \delta_5 \log(PROD_{jt}) + \delta_6 \log(DIST_{ij})$$
$$+ \delta_7 \log(KI_{it}) + \delta_8 SIM + \delta_9 RLF + \delta_{10} EMU$$
$$+ \delta_{11} EAST + \delta_{12} BORDER + \delta_{13} CEEC + \sum_{r=1993}^{2005} \delta_r t_r. \quad (5.7)$$

The dependent variable $EX_{ijt}$ in (5.4) represents country $i$'s exports to country $j$ for time period $t$ with an analogous notation for outward FDI ($FDIout_{ijt}$) in (5.5). The sub-indices for imports ($IM_{ijt}$) and inward FDI ($FDIin_{ijt}$) in (5.6) and (5.7) respectively, denote trade/FDI activity to $i$ from $j$ in period $t$. The use of time effects $t_r$ is motivated by findings in Baldwin and Taglioni (2006). The authors show that an exclusion of such time effects may result in significant misspecifications, given the fact that it is often impossible to obtain trade- or FDI-specific price data. Moreover, time effects allow us controlling for business cycle effects over the sample period. The other variables are defined as follows:

- $GDP$ = Gross domestic product in $i$ and $j$ respectively
- $POP$ = Population in $i$ and $j$
- $PROD$ = Labor productivity in $i$ and $j$
- $DIST$ = Geographical distance between state/national capitals
- $SIM$ = Similarity index defined as: $\log(1 - (\frac{GDP_{i,t}}{GDP_{i,t}+GDP_{j,t}})^2 - (\frac{GDP_{j,t}}{GDP_{i,t}+GDP_{j,t}})^2)$
- $RLF$ = Relative factor endowments in $i$ and $j$ defined as: $\log |(\frac{GDP_{i,t}}{POP_{i,t}}) - (\frac{GDP_{j,t}}{POP_{j,t}})|$
- $WAGE$ = Wage compensation per employee in $i$ and $j$
- $FDIopen$ = FDI openness in $j$ as share of total inward FDI relative to GDP
- $K$ = Total capital stock in $i$ and $j$
- $KI$ = Capital Intensity defined as Capital Stock per population in $i$
- $EMU$ = EMU membership dummy for $i$ and $j$
- $EAST$ = East German state dummy for $i$
- $BORDER$ = Border region dummy between $i$ and $j$
- $CEEC$ = Central and Eastern European country dummy for $j$

We can classify the set of control variables as either being time-varying or time-fixed. Time varying explanatory variables for the trade equations (both import & export flows) used throughout this analysis include *GDP* for home region and foreign country, population at home and abroad (*POP*), as well as variables, measuring the relative share of inter-industry trade (or vertical vs. horizontal FDI, respectively) based on indices of the similarity of economic size (*SIM*) and relative factor endowments (*RLF*).[9] The variable *SIM* captures the relative size of two countries in terms of *GDP*, assuming that we can model each German state as an individual small open economy (SOE). The variable takes values between zero (absolute divergence) and 0.5 (equal country size). *RLF* captures differences in terms of relative factor endowments, where we assume that these endowments are closely linked to per-capita GDP as a proxy for the former. The *RLF* variable takes a minimum of zero for equal factor endowments in the two regions. Based on recent findings in NTT models, we also test the effect of home and host country labor productivity (defined as GDP per total employment) on trade. We finally specify a (one) time-varying dummy to check for trade/FDI-creating effects of the EMU starting from 1999.

The economic interpretation of the time-varying variables is as follows: For the export equation (and imports vice versa) GDP levels at home and abroad are expected to be positively correlated with the level of exports (imports) reflecting the theoretical argument that the supply and demand for differentiated varieties increases with absolute higher income values. A similar connection can also be established if we substitute absolute income levels by per capita GDP in $i$ and $j$ as a proxy for welfare levels. The effect of population is not that clear cut. The most prominent interpretation is offered by Baldwin (1994) that both home and foreign country population levels are negatively related to trade, since larger countries tend to be more self-sufficient in terms of production and resource endowment. An alternative interpretation is that a positive impact of exporter population on trade indicates labor intensive good exports, while a negative one stands for capital intensive export dominance (see e.g. Serlenga and Shin 2007). In this line of argumentation, a positive correlation of foreign population and trade may indicate exports in necessity goods (likewise a negative one for luxury goods). Next to GDP or GDP per capita level we may also consider productivity measures at home and abroad. With respect to home (foreign) country productivity, we expect a positive influence on exports (imports) inspired by recent theoretical findings that more productive firms on average tend to have a higher degree of internationalization. *SIM* may serve as an indicator for the relative share of intra-industry trade. That is, the more similar countries are in terms of GDP, the higher will be the share of intra-industry trade. The interpretation of *RLF* is in similar veins (but of opposite coefficient sign). For increasing differences in factor endowments, we expect a rise in the relative share of inter-industry trade. For the EMU dummy we expect that the creation of the monetary unit has induced positive trade/FDI effects for its member states.

[9] In specifying the latter variables, we follow Egger (2001) and Serlenga and Shin (2007).

We use roughly the same set of time-varying variables for the gravity models of FDI (both inward and outward), and as Brenton et al. (1999) point out, the economic interpretation of the explanatory variables is very similar: As in the case of trade, FDI is expected to be positively related to the level of income at home and abroad as a proxy for a large domestic market, and negatively to population indicating that large population sized countries are expected to be more self-sufficient in terms of investment. An alternative interpretation would be that a positive correlation of FDI with a country's population indicates an FDI engagement of vertical type, since population is expected to the more abundant production factor with a lower price for labor. For transition countries (such as East Germany and CEECs) one could also consider a different interpretation of the population coefficient. Here the population level may capture the market potential effect of FDI much better than GDP related variables, reflecting the underlying hypothesis that the latter variables are still below their long-run trends alongside the catching-up process. Hence, population levels as a proxy for the market potential effect are assumed to be positively correlated with FDI activity. As for trade, we also include the variables *SIM* and *RLF* in the FDI equations as a potential indicator of the bilateral share of horizontal or vertical investment activities. Thereby, two similar countries (in terms of absolute GDP levels and/or factor endowments) are expected to engage more in horizontal than vertical FDI.

For the FDI models, we additionally augment the vector of time-varying variables by further endowment based variables derived from the NTT (see e.g. Borrmann et al. 2005). We include labor force specific skill variables and factor prices in the host country such as aggregate wage levels as well as FDI agglomeration forces proxied by the degree of FDI openness of the host country (e.g. defined as total inward FDI stock relative to GDP or alternatively the total per capita capital stock of the host country). We expect that agglomeration forces are typically positively related to the FDI activity. The effect of the wage level in the host country is a priori not clear. If vertical FDI activities are the dominant driving force, it should turn negative; for a dominance of horizontal FDI, a positive relationship between the wage level and FDI activity could also be true (indicating the need for a qualified workforce in foreign affiliate production and sales).

The set of time-invariant variables (both in the trade and FDI equations) includes geographic distance as proxy for transportation costs in the case of trade or fixed plant set-up and monitoring costs in the case of FDI. The role of distance has become one of the major research topics in trade theory, while typically a negative influence on both variables is assumed in the gravity model literature (see e.g. Markusen and Maskus 1999).[10] We further specify a dummy variable for differences in the export/FDI behavior of the East German states to capture historical and/or structural

[10] However, Egger and Pfaffermayr (2004) argue that although distance can be regarded as an obstacle to both trade and FDI, the two variables still may be seen as complements (rather than substitutes) with respect to this proxy for trade costs depending on the relative importance of plant set-up costs versus pure trade costs. Trade theory suggests that firms will tend to engage in FDI at the costs of trade as transport costs (proxied by distance) rise. More distant markets will tend to be served by overseas investments in firm affiliates rather than by exporting. Their

differences between the two German macro regions. Based on earlier research, we test the hypothesis whether the East German firms are still below their trade and investment potential.[11] We also test for neighboring (border) effects and measure the deviation of trade and FDI from German regions to CEECs compared to the core of the EU15 member states.[12]

Generally, neighboring effects are assumed to have a positive impact on trade and FDI due to historical, cultural and personal ties between the trading and investment partners. The expectations about the trade and FDI volume of German regions with the CEECs is not that clear a priori. For bilateral trade, several studies have revealed that German trade with the CEECs has increased rapidly after the transformation of these countries towards market economies in the early 1990s and that trade volumes now are already above their potential (relative to a normal trade level derived from the gravity model's determining factors) so that the dummy coefficient for trade is expected to be positive in particular for exports from Germany to the CEECs.[13] With respect to the FDI stock, it is questionable whether the short time span after the transformation to market economies is sufficient to build up a normal FDI stock (in the sense of the gravity model estimates), we thus expect a negative sign for the dummy variable coefficient with respect to outward FDI. The same logic applies for inward FDI. A summary of theoretically motivated coefficient signs for the gravity equations is given in Table 5.1.

## 5.3 Theory and Empirics of Trade-FDI Linkages

This section serves to give a short overview of recent theoretical and empirical contributions in analyzing trade-FDI linkages.[14] One basic observation is that the theoretical literature is rather inconclusive on that point since both type of interaction channels—either favoring a complementary or substitutive relations among the variables—can be found. The Heckscher–Ohlin (H–O) model with perfectly competitive product markets and no transportation costs as the standard workhorse model of traditional trade theory, for instance, explains trade between two countries mainly on differences in factor endowments. In the absence of factor mobility

---

hypothesis thus gives rise to a further proposal on how the estimate gravity models of trade and FDI properly, namely in an adequate simultaneous equations specification that explicitly accounts for the common determinants.

[11] See Alecke et al. (2003).

[12] The CEEC aggregate includes Hungary, Poland, the Czech Republic, Slovakia, Slovenia, Estonia, Latvia, Lithuania, Romania and Bulgaria.

[13] See e.g. Collins and Rodrik (1991), Wang and Winters (1991), Hamilton and Winters (1992), Baldwin (1994), Schumacher and Trübswetter (2000), Buch and Piazolo (2000), Jakab et al. (2001), Caetano et al. (2002) as well as Caetano and Galleg (2003).

[14] Markusen (1995), Jungmittag (1995), Zarotiadis and Mylonidis (2005) and Blanchard et al. (2008) among others provide detailed surveys of recent theoretical contributions.

**Table 5.1** Theoretically expected variable coefficients in the trade and FDI gravity equations

| Variable | Code | Trade eqs. | FDI eqs. | Expected coef. sign |
|---|---|---|---|---|
| Gross domestic product in $i/j$ | *GDP* (or $\frac{GDP}{POP}$) | X | X | (+) Trade/FDI activity increases with absolute higher income or welfare levels respectively (induced by higher supply and demand for differentiated varieties) |
| Population in $i/j$ | *POP* | X | X | (+/−) with − = Self-sufficiency in production (resource endowments); alternatively Trade: + = Δ share of labor intensive trade; FDI: + = market potential theory of FDI |
| Similarity index of $ij$ | *SIM* | X | X | (+/−) Trade: + = Δ share of intraindustry trade; FDI: + = Δ share of horizontal FDI |
| Relative factor endowments of $ij$ | *RLF* | X | X | (+/−) Trade: + = Δ share of interindustry trade; FDI + = Δ share of vertical FDI |
| Labor productivity in $i/j$ | *PROD* | X | X | (+) New Trade Theory: More productive firms on average higher degree of internationalization (expected to be higher for FDI than Trade) |
| Euro area dummy | *EMU* | X | X | (+) Trade/FDI creating effect of single currency |
| Wage level in $j$ | *WAGE* | | X | (−) Indicator for vertical cost oriented FDI engagement (only in outward FDI equation) |
| FDI openness in $j$ | *FDIopen* | | X | (+) Proxy for agglomeration forces at work (only in outward FDI equation) |
| Capital stock in $j$ | *K* | | X | (+/−) with + = Agglomeration forces or − = Neoclassical view (H–O) of higher expected return for relatively scare production factor (only in outward FDI equation) |
| Capital intensity in $i$ | *KI* | | X | (+/−) with + = Agglomeration forces or − = Neoclassical view (H–O) of higher expected return for relatively scare production factor (only in inward FDI equation) |
| Geographical distance of $ij$ | *Dist* | X | X | (+/−) Trade: − = Transportation costs as obstacles to trade; FDI: + = FDI as alternative to trade for increasing distances, alternatively: − = Increasing monitoring costs over longer distance, increasing cultural differences etc. |
| East German State dummy | *East* | X | X | (+/−) A-priori unknown (possibly: − = Negative historical path dependency in East German internationalization process) |
| CEE Country dummy | *Ceec* | X | X | (+/−) A-priori unknown (possibly: − = Negative historical path dependency in CEEC internationalization process) |
| Common Border dummy | *Border* | X | X | (+) Positive neighboring effect on trade/FDI due to historical, cultural and personal ties |

(FDI), international trade serves to equalize factor prices across countries. However, if factor mobility increases, the differences in endowments diminish and trade volumes tend to decrease. Surveying recent theoretical contributions, Markusen (1995) shows that the substitutive H–O model predictions can also be extended to the case of imperfect competition. A prominent approach of the latter type is the so-called proximity-concentration trade-off explored by Brainard (1993, 1997). Here, under the assumption of non-zero trade costs, the extent to which firms decide to engage in trade rather than foreign sales (FDI) depends crucially on the relative benefits of being close to the targeted market versus concentrating production in one location, which is associated with the exploitation of economies of scale.

On the contrary, recent contributions also derive complementaries between trade and FDI. A starting point is the General Equilibrium model of Helpman (1984), which models MNEs as vertically integrated firms in a monopolistic competition environment with their choice of location for (intermediate) production being driven by relative factor costs and resource endowments. In this set-up, FDI is more likely to create (inter-industry) trade rather than replace it. Consequently, from a vertically integrated modelling perspective, trade and FDI are complementary with respect to differences in factor endowments. Starting from a critical reflection of the proximity-concentration trade-off literature, Baldwin and Ottaviano (2001) show that complementary and substitutive elements in trade-FDI activity may coexist. In their model, multi-product final-good producing firms simultaneously engage in intra-industry trade and FDI based on the idea that obstacles to trade generate a natural incentive for multi-product firms to do so. In the model, non-zero trade costs shift production location to foreign affiliates so that, as a result, FDI displaces some exports (as standard trade theory result). However, it may also enhance trade via reverse imports of final goods since products in the model are differentiated. One of the advantages of the model is that the parallelism between the pattern of trade and investment is at the core of the model's driving mechanism. For our empirical analysis of German trade/FDI activity within the EU27, the model may be seen as especially relevant, since it is explicitly designed to explain the behavior of European MNEs and track the specific European trade-FDI pattern/nexus—with Europe being modelled as a rather closed trading area.

There are also various approaches aiming to pin down the trade-FDI-nexus empirically. Though on average there is a general tendency to reveal complementary linkages, the empirical literature also gives heterogeneous answers to this question. As Aizenman and Noy (2006) point out, important aspects to account for in the empirical set-up are to closely interpret the estimation result in light of the chosen country, industry sample and time period under observation. That is, for example, with respect to positive trade-FDI linkages much more empirical support is found in the context of developing rather than developed countries (see e.g. Tadesse and Ryan 2004). Another sensitive aspect is the sample period. As Pain and Wakelin (1998) point out, the nature of the trade-FDI linkage may change over time e.g. depending on the maturity of the investments and the accumulation of investments over time in terms of a country's stage of internationalization activity.

Empirical approaches may be broadly classified into macro and micro (firm-level) studies. The latter are typically characterized by a detailed sectoral disaggregation and accounts for firm heterogeneity, whereas the former analysis puts trade and FDI flows in its macroeconomic context. Aggregate data are predominantly estimated in a gravity model framework, mainly focusing on the link between exports and outward FDI. Selected results of the empirical literature for industrialized countries are as follows: For US data, Lipsey and Weiss (1981, 1984) find a positive coefficient in regressing US outward FDI stocks on exports. Subsequently Brainard (1997), Graham (1999), Clausing (2000), Egger and Pfaffermayr (2004) as well as Fontagne and Pajot (1997) support this complementary view. For the UK Zarotiadis and Mylonidis (2005) find positive ties between trade and FDI based on inward FDI stocks as well as both export and import data. In the case of Japan the picture is rather different with the majority of studies revealing substitutive linkages: A negative export-outward FDI nexus is e.g. reported in Ma et al. (2000) and Bayoumi and Lipworth (1997). Only Nakamura and Oyama (1998) find trade expansion effects of outward FDI. For other country pairs (including a macro-sectoral disaggregation) studies such as Bloningen (2001) for USA–Japanese trade and FDI relations as well as Goldberg and Klein (1999) for the USA and South American countries reveal mixed evidence with both complementary and substitutive elements depending on the chosen country and sector under considerations. Among the few studies using (West) German data, Jungmittag (1995) and Egger and Pfaffermayr (2004) identify substitutive relationships—however solely focusing on exports and outward FDI stock. We also add imports and inward FDI to the analysis.

## 5.4 Data and Stylized Facts

We use a panel data set for 16 German states (*Bundesländer*) and the EU27 member countries, which gives a total of 368 country pairs (16 states $\times$ 23 countries).[15] Our database covers a time period of 13 years (1993–2005). Due to data limitations, we have to cope with an unbalanced panel. Import and export data is balanced for the whole sample. In the FDI equation we distinguish between zero FDI stock and not reported values. The latter are handled as missing data while we substitute zero trade flows by a small constant while using log-linear gravity models. For an overview of different methods of dealing with zero trade flows in the gravity model context see e.g. Linders and de Groot (2006). Though Coe and Tamirisa (2002) show that the results may differ significantly when excluding zero flows in the log-linear specification, our results remain rather stable when using different proxies for these zeros. A complete list of variables and data sources is given in Table 5.2.

Before we turn to the specification of the empirical model, we highlight some stylized facts of German trade and FDI patterns both from an aggregated as well as

[15] We exclude Malta and Cyprus due to their specific characteristics as island economies. Further, we treat Belgium and Luxembourg as one single economy mainly due to the limited accessibility of statistical data.

**Table 5.2** Data description and source

| Variable | Description | Source |
|---|---|---|
| $EX_{ijt}$ | Export volume, nominal values, in Mio. € | Destatis (2008) |
| $IM_{ijt}$ | Import volume, nominal values, in Mio. € | Destatis (2008) |
| $FDIout_{ijt}$ | Outward FDI stock, nominal values, in Mio. € | Deutsche Bundesbank (2008) |
| $FDIin_{ijt}$ | Inward FDI stock, nominal values, in Mio. € | Deutsche Bundesbank (2008) |
| $GDP_{it}$ | Gross domestic product, nominal values, in Mio. € | VGR der Länder (VGRdL 2008) |
| $GDP_{jt}$ | Gross domestic product, nominal values, in Mio. € | Eurostat (2008) |
| $POP_{it}$ | Population, in 1000 | VGRdL (2008) |
| $POP_{jt}$ | Population, in 1000 | Groningen Growth & Development center (GGDC 2008) |
| $SIM_{ijt}$ | $SIM = \log\left(1 - \left(\frac{GDP_{it}}{GDP_{it}+GDP_{jt}}\right)^2 - \left(\frac{GDP_{jt}}{GDP_{it}+GDP_{jt}}\right)^2\right)$ | see above |
| $RLF_{ijt}$ | $RLF = \log\left\lvert\left(\frac{GDP_{it}}{POP_{it}}\right) - \left(\frac{GDP_{jt}}{POP_{jt}}\right)\right\rvert$ | see above |
| $EMP_{it}$ | Employment, in 1000 | VGRdL (2008) |
| $EMP_{jt}$ | Employment, in 1000 | EU Commission (2008) |
| $PROD_{it}$ | $Prod_{it} = \left(\frac{GDP_{it}}{EMP_{it}}\right)$ | see above |
| $PROD_{jt}$ | $Prod_{jt} = \left(\frac{GDP_{jt}}{EMP_{jt}}\right)$ | see above |
| $K_{it}$ | Capital stock, nominal, in Mio. € | VGRdL (2008) |
| $K_{jt}$ | Capital stock derived from GFCF via perpetual inventory method, nominal, in Mio. € | GFCF data from Eurostat (2008) |
| $KI_{it}$ | $KI_{it} = \left(\frac{K_{it}}{POP_{it}}\right)$ | see above |
| $FDIopen_{jt}$ | $FDIopen_{jt} = \left(\frac{Total\ inward\ FDI_{jt}}{GDP_{jt}}\right)$ | FDI: (2008), GDP: see above |
| $WAGE_{it}$ | Wage compensation per employee, nominal, in 1000 | VGRdL (2008) |
| $WAGE_{jt}$ | Wage compensation per employee, nominal, in 1000 | EU Commission (2008) |
| $DIST_{ij}$ | Distance between state capital for Germany and national capital for the EU27 countries, in km | Calculation based on coordinates, calculation tool obtained from www.koordinaten.de |
| $EMU$ | (0, 1)-dummy variable for EMU members since 1999 | |
| $EAST$ | (0, 1)-dummy variable for the East German states | |
| $CEEC$ | (0, 1)-dummy variable for the Central and Eastern European countries | |
| $BORDER$ | (0, 1)-dummy variable for country pairs with a common border | |
| $t_{1993}$–$t_{2005}$ | Time effects for the years 1993–2005 | |

a regional perspective. One of the main characteristics of the German economy is its relative strong openness to international trade and FDI. In 2005 German exports accounted for approximately 9.5% of total worldwide merchandise flows—making Germany the world's leading exporting nation worldwide ahead of the USA (8.9%), China (7.5%) and Japan (5.9%). Taking a closer look at the bilateral trade pattern with Germany's major trading partners, for import flows six out of the ten major partners come from the EU27 and for exports these are even eight out of ten. The share of German-EU27 trade relative to worldwide trade is 67.2% (for the average of 1993–2005) and for imports it is almost equally high (64.8%). Compared to exports the EU27-wide outward FDI share is somewhat lower (51.9% between 1993–2005) but still amounts to a significant part.[16] The percentage share of the inward FDI stock from EU countries for this period is extremely high in the case of Germany (73.8% relative to total inward FDI).

Looking at German regional trade and FDI intensities (defined as regional trade/FDI per regional GDP), Table 5.3 reports regional differences relative to the German average (where the latter is normalized to one). States with the highest total export intensity are Bremen (1.83 for 2000–2005), Saarland (1.47) and Baden-Württemberg (1.36). The figures are roughly similar for total as well as intra-EU exports. One major exception is the Saarland which has a significantly higher intra-EU trade intensity (1.91) compared to the total trade intensity (1.47). Since Saarland has a common border with France (and strong historical and cultural ties), this may be seen as an indication of a positive trade effect of a common border and close distance ties to EU trading partners, which are typically tested in a gravity model context. The most import intensive regions apart from the city states Bremen and Hamburg are Hessen (1.12 for total imports between 2000 and 2005), North Rhine-Westphalia (1.12) and Saarland (1.45). Examining the differences between the two West and East German macro regions, Table 5.3 shows that the East German states trade roughly half as much as the German average indicating that the East German states are still less involved in international trade compared to their Western counterparts. Figure 5.1 displays the results graphically.

With respect to the FDI intensities Table 5.3 shows that the southern states Hessen (2.32 for the period 2000 to 2005), Baden-Württemberg (1.33) and Bavaria (1.15) have the highest outward FDI activity after adjusting for absolute GDP levels. For the five East German states (Brandenburg, Mecklenburg-Vorpommern, Saxony, Saxony-Anhalt and Thuringia), the outward FDI activity is extremely low (0.06 for total and 0.04 for intra-EU FDI stocks). Looking at inward FDI the West–East gap is somewhat smaller, mirroring the broad picture that the Eastern states throughout their economic transition process are able to act as a host country for FDI, but with little options for East German firms to actively invest abroad. The (macro) regional differences for German trade-FDI activity are also summarized graphically in Fig. 5.1. The regional perspective of German state export and FDI activity shows

[16]The remainder part of Germany's outward FDI stock is mainly directed to the US (29.6% in 2005).

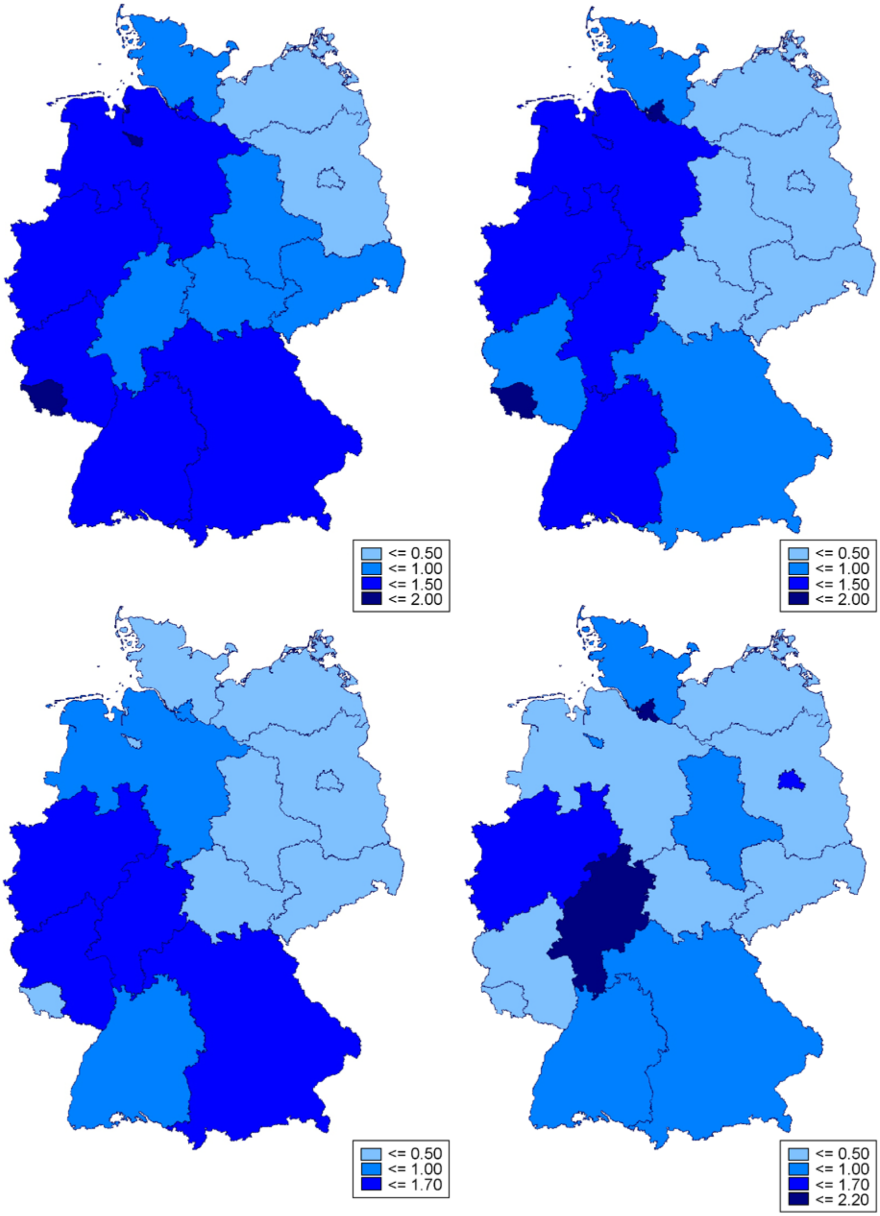


**Fig. 5.1** Regional trade and FDI intensities within the EU27 for average 2000–2005 (with *upper left*: exports, *upper right*: imports, *lower left*: outward FDI, *lower right*: inward FDI). *Source*: See Table 5.3

that we detect strong regional difference for which we have to account when setting up a model that includes economic and geographic variables in explaining the German export and FDI performance.

**Table 5.3** Relative export, import, outward and inward FDI intensity of German states compared to the national average (Germany = 1)

| | Export intensity | | | | Import intensity | | | |
|---|---|---|---|---|---|---|---|---|
| | Av. 1993–99 | | Av. 2000–05 | | Av. 1993–99 | | Av. 2000–05 | |
| | World | EU27 | World | EU27 | World | EU27 | World | EU27 |
| BW | 1.41 | 1.25 | 1.36 | 1.23 | 1.00 | 0.99 | 1.09 | 1.08 |
| BAY | 1.09 | 1.07 | 1.10 | 1.05 | 0.96 | 0.98 | 0.95 | 0.95 |
| BER | 0.46 | 0.42 | 0.46 | 0.42 | 0.31 | 0.35 | 0.33 | 0.33 |
| BRA | 0.31 | 0.35 | 0.42 | 0.44 | 0.46 | 0.44 | 0.54 | 0.42 |
| BRE | 1.97 | 1.70 | 1.83 | 1.64 | 2.62 | 1.45 | 1.87 | 1.36 |
| HH | 0.86 | 0.86 | 1.10 | 1.12 | 2.20 | 1.50 | 2.15 | 1.58 |
| HES | 0.82 | 0.82 | 0.71 | 0.69 | 1.27 | 1.19 | 1.12 | 1.08 |
| MV | 0.27 | 0.22 | 0.34 | 0.33 | 0.24 | 0.34 | 0.28 | 0.33 |
| NIE | 1.06 | 1.13 | 1.09 | 1.18 | 0.91 | 0.95 | 1.06 | 1.05 |
| NRW | 1.10 | 1.17 | 1.03 | 1.10 | 1.18 | 1.26 | 1.12 | 1.21 |
| RHP | 1.26 | 1.31 | 1.18 | 1.22 | 0.93 | 1.04 | 0.81 | 0.97 |
| SAAR | 1.43 | 1.76 | 1.47 | 1.91 | 1.25 | 1.64 | 1.45 | 1.97 |
| SACH | 0.36 | 0.41 | 0.68 | 0.61 | 0.33 | 0.44 | 0.43 | 0.48 |
| ST | 0.32 | 0.34 | 0.45 | 0.53 | 0.29 | 0.33 | 0.44 | 0.37 |
| SH | 0.69 | 0.66 | 0.73 | 0.74 | 0.75 | 0.82 | 0.82 | 0.90 |
| TH | 0.37 | 0.39 | 0.54 | 0.58 | 0.33 | 0.41 | 0.43 | 0.45 |
| *Germany* | *1.00* | *1.00* | *1.00* | *1.00* | *1.00* | *1.00* | *1.00* | *1.00* |
| *East** | *0.33* | *0.36* | *0.52* | *0.52* | *0.34* | *0.40* | *0.43* | *0.43* |
| *West** | *1.11* | *1.11* | *1.09* | *1.09* | *1.12* | *1.11* | *1.11* | *1.11* |

| | Outward FDI intensity | | | | Inward FDI intensity | | | |
|---|---|---|---|---|---|---|---|---|
| | Av. 1993–99 | | Av. 2000–05 | | Av. 1993–99 | | Av. 2000–05 | |
| | World | EU27 | World | EU27 | World | EU27 | World | EU27 |
| BW | 1.24 | 0.97 | 1.33 | 0.89 | 0.90 | 0.87 | 0.77 | 0.70 |
| BAY | 1.29 | 1.41 | 1.15 | 1.44 | 0.67 | 0.68 | 0.90 | 0.96 |
| BER | 0.50 | 0.62 | 0.24 | 0.28 | 0.73 | 0.82 | 1.04 | 1.14 |
| BRA | 0.06 | 0.06 | 0.02 | 0.03 | 0.32 | 0.46 | 0.27 | 0.31 |
| BRE | 0.27 | 0.41 | 0.10 | 0.15 | 1.03 | 1.24 | 0.76 | 0.81 |
| HH | 1.08 | 1.33 | 0.67 | 0.80 | 2.00 | 2.02 | 1.89 | 2.15 |
| HES | 2.02 | 2.03 | 2.32 | 1.65 | 2.59 | 1.95 | 2.34 | 1.88 |
| MV | 0.12 | 0.03 | 0.03 | 0.04 | 0.39 | 0.37 | 0.37 | 0.29 |
| NIE | 0.77 | 0.84 | 0.62 | 0.76 | 0.59 | 0.61 | 0.50 | 0.45 |
| NRW | 0.99 | 1.00 | 1.16 | 1.34 | 1.21 | 1.29 | 1.29 | 1.44 |
| RHP | 1.25 | 1.21 | 1.04 | 1.32 | 0.56 | 0.73 | 0.50 | 0.50 |

(*continued on the next page*)

**Table 5.3** (Continued)

| | Outward FDI intensity | | | | Inward FDI intensity | | | |
|---|---|---|---|---|---|---|---|---|
| | Av. 1993–99 | | Av. 2000–05 | | Av. 1993–99 | | Av. 2000–05 | |
| | World | EU27 | World | EU27 | World | EU27 | World | EU27 |
| SAAR | 0.44 | 0.66 | 0.25 | 0.36 | 0.58 | 1.00 | 0.40 | 0.47 |
| SACH | 0.02 | 0.01 | 0.06 | 0.02 | 0.20 | 0.17 | 0.17 | 0.10 |
| ST | 0.11 | 0.00 | 0.01 | 0.00 | 0.97 | 1.70 | 0.59 | 0.78 |
| SH | 0.19 | 0.18 | 0.14 | 0.17 | 0.52 | 0.49 | 0.64 | 0.63 |
| TH | 0.06 | 0.06 | 0.06 | 0.15 | 0.23 | 0.35 | 0.23 | 0.15 |
| *Germany* | *1.00* | *1.00* | *1.00* | *1.00* | *1.00* | *1.00* | *1.00* | *1.00* |
| *East*[*] | *0.06* | *0.03* | *0.04* | *0.04* | *0.40* | *0.56* | *0.30* | *0.30* |
| *West*[*] | *1.15* | *1.15* | *1.16* | *1.16* | *1.09* | *1.07* | *1.09* | *1.09* |

*Note*: BW = Baden-Württemberg, BAY = Bavaria, BER = Berlin, BRA = Brandenburg, BRE = Bremen, HH = Hamburg, HES = Hessen, MV = Mecklenburg-Vorpommern, NIE = Lower Saxony, NRW = North Rhine-Westphalia, RHP = Rhineland-Palatine, SAAR = Saarland, SACH = Saxony, ST = Saxony-Anhalt, SH = Schleswig-Holstein, TH = Thuringia

*Source*: Data from Destatis (2008), Deutsche Bundesbank (2008), VGRdL (2008)

[*]East = East German states (excluding Berlin), West = West German states (excluding Berlin)

## 5.5 Econometric Specification and Estimation Results

### *5.5.1 Time Series Properties of the Variables*

With the gravity model literature having its root in cross-sectional studies little attention has been typically paid to the time-series properties of the variables even if the empirical application now predominantly has switched to panel data estimation (exceptions are e.g. Fidrmuc 2009; Zwinkels and Beugelsdijk 2010). While for the standard microeconometric panel data model with $N \to \infty$ and fixed $T$, the assumption of stationarity may be seen as justified, it becomes less evident for macro panels with an increasing time dimension. Since our data with $N = 353$ and maximum $T = 13$ is at the borderline between classical micro and macro panel data, we aim to explicitly account for the time-series properties in order to avoid the problem of spurious regression among non-stationary variables that are not cointegrated.

Different approaches have been proposed to test for unit roots in panel data. However, only few are directly applicable to unbalanced data without inducing a bias to the test results (see e.g. Baltagi 2008, for an overview). Here we rely on a Fisher-type testing approach which averages the $p$-values of unit root tests for each cross section $i$ as proposed by Maddala and Wu (1999) and Choi (2001). The null hypothesis of the test is that the series under observation is non-stationary. Fidrmuc (2009) alternatively proposes the CADF test from Pesaran (2007), which also works with unbalanced panel data. We use the CADF test to double check for those variables we do not reject the null hypothesis of a unit root in the series based on the Fisher-type test.

The results of the panel unit root tests for the variables in levels are given in Table 5.4. The results predominantly reject the null hypothesis of non-stationarity for

**Table 5.4** Fisher-type and Pesaran (2007) panel unit root tests for variables in levels

| Variables | $\chi^2$-statistic of Fisher-type test ($p$-val.) $H_0$: Series non-stationary | |
|---|---|---|
| | Constant without trend | Constant and time trend |
| $EX_{ijt}$ | 813.08*** (0.00) | 842.63*** (0.00) |
| $FDIout_{ijt}$ | 853.27*** (0.00) | 687.85*** (0.00) |
| $IM_{ijt}$ | 1099.67*** (0.00) | 821.67*** (0.00) |
| $FDIin_{ijt}$ | 602.89 (0.26) | 579.81 (0.51) |
| $GDP_{it}$ | 1412.13*** (0.00) | 1364.72*** (0.00) |
| $GDP_{jt}$ | 522.63 (0.96) | 772.73*** (0.00) |
| $POP_{it}$ | 2744.13*** (0.96) | 502.02 (0.99) |
| $POP_{jt}$ | 2171.32*** (0.00) | 1160.79*** (0.00) |
| $PROD_{it}$ | 1224.90*** (0.00) | 1669.38*** (0.00) |
| $PROD_{jt}$ | 413.19 (0.99) | 827.45*** (0.00) |
| $SIM_{ijt}$ | 783.17*** (0.00) | 1096.57*** (0.00) |
| $RLF_{ijt}$ | 565.87 (0.67) | 1012.69*** (0.00) |
| $WAGE_{jt}$ | 554.41 (0.78) | 759.67*** (0.00) |
| $FDIopen_{jt}$ | 628.54* (0.08) | 233.97 (0.99) |
| $K_{jt}$ | 2387.88*** (0.00) | 804.83*** (0.00) |
| $KI_{it}$ | 1609.78*** (0.00) | 1084.10*** (0.00) |

| Critical Vars. | $\overline{CADF}$ for Pesaran (2007) test ($p$-val.) $H_0$: Series non-stationary | |
|---|---|---|
| | Constant without trend | Constant and time trend |
| $FDIin_{ijt}$ | 22.11 | 21.62 |
| $GDP_{jt}$ | −3.75*** | 4.94 |
| $POP_{it}$ | −3.67*** | 3.85 |
| $PROD_{jt}$ | −4.36*** | 5.58 |
| $RLF_{ijt}$ | −9.68*** | −5.77*** |
| $WAGE_{jt}$ | −16.14*** | −3.44*** |
| $FDIopen_{jt}$ | −6.38*** | −0.29 |

*Note*: $p$-values are given in parenthesis. Critical values for the CADF test are taken from Pesaran (2003). These are for panel regression with $T = 15$, $N = 200$ including a regression constant but no trend: 1% (−2.16), 5% (−2.04), 10% (−1.98). For the test alternative with constant and time trend: 1% (−2.71), 5% (−2.57), 10% (−2.50)
*Denote statistical significance at the 10% level **Denote statistical significance at the 5% level
***Denote statistical significance at the 1% level

the variables in the dataset. However, both the Fisher-type unit root test as well as Pasaran's CADF test detect some cases which indicate non-stationarity of the time series. Since there is some heterogeneity with respect to the chosen test statistic, we are cautious in using the results unambiguously in favor of stationarity and additionally perform a residual-based unit root test for panel cointegration in the spirit of Kao (1999) on our final model specification to avoid the risk of running spurious regressions.

### 5.5.2 Econometric Specification

In estimating the system in (5.4)–(5.7) we carefully account for the trade-off between the likely increase in estimation efficiency based on a full information system approach (if we observe a significant correlation of the residuals from a single equation estimation of the respective gravity models) and the additional complexity brought into the system, which in turn may translate into increasingly biased results if the estimation error of one equation is transmitted to all other equations. The use of simultaneous equations models with panel data is not that common. However, Cornwell et al. (1992), Baltagi (2008), Baltagi and Chang (2000), Prucha (1984), Krishnakumar (1988) as well as Park (2005), among others, discuss both fixed effects and random effects panel data estimators in a system manner where right hand side endogeneity matters. The goal is to apply both IV and non-IV approaches to our simultaneous equation approach for the trade/FDI system. IV estimation thereby builds on the Hausman–Taylor (1981) model as the standard estimator in the field, while the non-IV alternative centers around a two-step estimator based on the Fixed Effects model, which has shown a good performance both in Monte Carlo simulations and empirical applications to gravity type models recently.

The Hausman–Taylor (1981) model may be seen as a hybrid version of the Fixed Effects (FEM) and Random Effects (REM) model. The idea of the Hausman–Taylor (HT) estimator is to derive consistent instruments from internal data transformations to cope with endogeneity, but still to avoid the strong all-or-nothing assumptions of the FEM and REM in terms of residual correlation of the right hand side regressors respectively. The Hausman–Taylor model therefore splits both the vectors of time-varying and time-fixed variables into two sub-vectors classifying the variables as either being correlated or uncorrelated with the unobservable individual effects. This classification scheme is then used to derive consistent instruments for model estimation.

We use the HT setup for estimating a 3SLS-GMM estimator, which has the advantage over standard 3SLS estimation because it allows the use of different instruments in subsequent equations of the system, while standard 3SLS assumes the same IV-set applies to every equation in the system. The latter assumption may be somewhat problematic in our case, since we have found that different instruments are valid for subsequent model equations based on a series of Hansen (1982)/Sargan (1958) overidentification tests for single equation benchmark models.[17] For convenience and in line with the mainstream literature on the Hausman–Taylor model, we assume that the variance-covariance (VCV) matrix of the error terms takes the random effects form.

[17] Results are not reported here, but can obtained upon request.

As alternative to the Hausman–Taylor IV estimator, we further apply a non-IV two-step modelling approach, which basically builds on the Fixed Effects Model (FEM) but also allows us to quantify the effect of time-fixed variables, which are wiped out by the within-type data transformation in the standard FEM. To avoid this problem, the two-step approach estimates the coefficient vector of the time-varying variables by FEM in a first step and then applies pooled OLS (POLS) in a second step to obtain a vector of coefficients for these variables that involves a regression of the first step group mean residuals (as a proxy for the unobserved individual effects) against the vector of time-fixed variables. Since this second step includes a generated regressand we have to adjust the standard errors. Due to the decomposition of the vector of fixed effects Plümper and Tröger (2007) label the estimator as Fixed Effects Vector Decomposition (FEVD).[18]

One advantage of the non-IV specification compared to the Hausman–Taylor approach is that no arbitrary ex-ante selection of consistent moment conditions (IVs) is necessary, and the approach avoids the risk of running into the weak-instrument problem, which may well apply to the former approach and result in a substantial finite sample bias. The FEVD-type two-step estimator has recently been applied in a variety of empirical contributions; especially for gravity type models (see e.g. Belke and Spies 2008, as well as Caporale et al. 2008). Small sample based Monte Carlo simulation experiments have confirmed the overall good empirical performance of this non-IV approach, which is found to be superior relative to the HT estimator especially in terms of getting the time-fixed variable coefficients right (see e.g. Plümper and Tröger 2007; Mitze 2009).

In the context of the FEVD-type two-step estimator the adaptation to a system approach is rather straightforward. That is, for the FEM model, Cornwell et al. (1992) show that in the absence of any assumption about the individual effects, one cannot do better than apply any efficient system estimator to the within-type transformed model. Analogously, for POLS—which ignores individual heterogeneity—the model can be directly applied in a seemingly unrelated regression (SUR) framework adjusting for the system's residual VCV matrix of the system by GLS estimation. In line with the FEVD single equation approach by Plümper and Tröger (2007), we will label the newly proposed system extension throughout the remainder of our analysis as FEVD-SUR. To adjust standard errors (SE) in the second step, we choose bootstrapping techniques as discussed in Atkinson and Cornwell (2006). We apply the wild bootstrap procedure, which has shown a good empirical performance in a variety of Monte Carlo simulation experiments (see e.g. Davidson and Flachaire 2001; MacKinnon 2002, and Atkinson and Cornwell 2006).[19]

For both the IV and non-IV approach, we apply the same estimation strategy. We first estimate the individual equations of the system in (5.4)–(5.7) and test for the cross-equation correlation of residuals, which indicate the use of a full information

---

[18] The reader is referred to Appendix A for a detailed discussion of the estimation settings of the FEVD.

[19] Additional details on the specification of both estimators including the bootstrapping procedure for the FEVD-SUR are given in Appendix A.

approach. On the fly, this approach allows us to derive a measure of the underlying trade-FDI linkages for our sample of German regions based on the first step estimates of the system's residual VCV matrix as pointed out by Egger and Pfaffermayr (2004). In this logic, elements beside the main diagonal in the VCV matrix of the (composed) error term can be used as estimates for the underlying state-country pair trade and FDI linkages. A negative parameter indicates a substitutive relationship between the two analyzed variables after controlling for common and observed exogenous determinants. The test setup may be seen as a straightforward extension to the standard approach to test for trade-FDI linkages, which typically employ simple pairwise residual correlations in an auxiliary regression (e.g. Graham 1999; Brenton et al. 1999; Pantulu and Poon 2003; Africano and Magalhaes 2005, among others). We use Breusch–Pagan (1980) type LM tests corrected for unbalanced panel data sets according to Song and Jung (2001) and Baltagi and Song (2006) to check for the significance of the cross-equation residual correlation.[20]

### 5.5.3 Estimation Results

Table 5.5 plots the results for the Hausman–Taylor 3SLS-GMM estimator and Table 5.6 reports the FEVD-SUR findings. The $R^2$ shows that both estimates are quite

**Table 5.5** 3SLS-GMM estimation results for Hausman–Taylor model

| Dep. variable | HT-3SLS-GMM | | | |
|---|---|---|---|---|
| | Exports | FDI out | Imports | FDI in |
| $Log(GDP_i)$ | 0.94 | 5.11*** | 1.23** | 2.58*** |
| | (0.650) | (1.777) | (0.503) | (0.996) |
| $Log(GDP_j)$ | 0.12 | 0.93*** | 2.65*** | 5.56*** |
| | (0.948) | (0.242) | (0.855) | (1.085) |
| $Log(POP_i)$ | −1.55** | −3.35** | −0.42 | 1.35* |
| | (0.769) | (1.688) | (0.533) | (0.781) |
| $Log(POP_j)$ | 0.58*** | 2.31*** | −1.88** | −6.49*** |
| | (0.146) | (0.404) | (0.858) | (1.177) |
| $Log(PROD_i)$ | 2.01*** | −3.92** | | |
| | (0.638) | (1.904) | | |
| $Log(PROD_j)$ | | | −2.52*** | −5.50*** |
| | | | (0.821) | (1.092) |
| $Log(DIST_{ij})$ | −1.23*** | −3.21*** | −1.53*** | −2.88*** |
| | (0.366) | (0.497) | (0.311) | (0.904) |
| $Log(WAGE_j)$ | | 0.13 | | |
| | | (0.271) | | |

(*continued on the next page*)

[20]Further details on the specification of the test statistic are given in Appendix B.

**Table 5.5** (Continued)

| Dep. variable | HT-3SLS-GMM | | | |
|---|---|---|---|---|
| | Exports | FDI out | Imports | FDI in |
| $Log(FDIopen_j)$ | | 0.49*** | | |
| | | (0.131) | | |
| $Log(KF_j)$ | | −0.95*** | | |
| | | (0.344) | | |
| $Log(\frac{KBL_i}{POP_i})$ | | | | −2.26*** |
| | | | | (0.678) |
| *SIM* | −0.37*** | 1.24*** | −0.69*** | −0.52* |
| | (0.102) | (0.349) | (0.248) | (0.317) |
| *RLF* | 0.01 | 0.01 | 0.07** | −0.06 |
| | (0.010) | (0.034) | (0.034) | (0.041) |
| *EMU* | 0.20*** | −0.51*** | 0.04 | 0.57*** |
| | (0.041) | (0.143) | (0.067) | (0.164) |
| *EAST* | −0.79*** | −2.98*** | 0.36 | 2.12*** |
| | (0.203) | (0.475) | (0.282) | (0.522) |
| *BORDER* | 0.73 | −1.22* | 0.29 | −1.72 |
| | (0.590) | (0.691) | (0.430) | (1.399) |
| *CEEC* | −0.48* | −3.15*** | 0.15 | −3.99*** |
| | (0.285) | (0.533) | (0.359) | (0.629) |
| Time effects | Yes | Yes | Yes | Yes |
| ($p$-value of Wald test) | (0.00) | (0.00) | (0.00) | (0.00) |
| No. of system observation | 10660 | | | |
| No. of obs. per equation | 2665 | 2665 | 2665 | 2665 |
| No. of groups per equation | 353 | 353 | 353 | 353 |
| KP weak ident. $F$-test | 38.64 | 85.12 | 147.98 | 21.98 |
| Staiger–Stock rule ($F \geq 10$) | passed | passed | passed | passed |
| Hansen/Sargan overid. | 8.67 (3) | 9.98 (4) | 8.53 (5) | 42.86 (3) |
| ($p$-value) | (0.04) | (0.04) | (0.12) | (0.00) |
| $\|m\|$-stat. 3SLS/2SLS | 0.01 | 28.56 | 42.26 | 36.54 |
| ($p$-value) | (0.99) | (0.43) | (0.01) | (0.08) |
| Resid. based ADF test | 766.4*** | 1113.5*** | 1579.9*** | 1327.0*** |
| ($p$-value) | (0.00) | (0.00) | (0.00) | (0.00) |
| $R^2$ | 0.69 | 0.66 | 0.42 | 0.59 |

*Note*: Standard errors are robust to heteroscedasticity and clustering on bilateral pairs. Variable classification: $X1 = [GDP^1_{jt}, POP^1_{jt}, PROD^1_{jt}, POP^2_{jt}, POP^2_{it}, PROD^2_{jt}, WAGE^2_{jt}, KF^2_{jt}, GDP^3_{it}, GDP^3_{jt}, POP^3_{jt}, POP^3_{it}, PROD^3_{jt}, RLF^3_{ijt}, POP^4_{jt}, PROD^4_{jt}, KBLC^4_{it}, RLF^4_{ijt}]$ and $Z2 = [DIST^1_{ij}, DIST^2_{ij}, DIST^3_{ij}]$, where high level indices label the equation number as 1 = export, 2 = outward FDI, 3 = imports, 4 = inward FDI. Endogeneity of $Z2$ variables is tested based on the $C$-statistic

*Denote statistical significance at the 10% level **Denote statistical significance at the 5% level

***Denote statistical significance at the 1% level

**Table 5.6** FEVD-SUR estimation results

| Dep. variable | FEVD-SUR | | | |
|---|---|---|---|---|
| | Exports | FDI out | Imports | FDI in |
| $Log(GDP_i)$ | 0.62* | 4.50*** | 1.56*** | 1.57*** |
| | (0.356) | (1.263) | (0.215) | (0.572) |
| $Log(GDP_j)$ | 0.13** | −0.85 | 1.35*** | 4.91*** |
| | (0.056) | (0.552) | (0.177) | (0.429) |
| $Log(POP_i)$ | −1.57*** | −1.30 | −0.70 | 6.79*** |
| | (0.527) | (1.847) | (0.455) | (1.314) |
| $Log(POP_j)$ | 2.17*** | −0.52 | 2.89*** | −0.70 |
| | (0.410) | (1.440) | (0.548) | (1.345) |
| $Log(PROD_i)$ | 2.16*** | −4.34*** | | |
| | (0.362) | (1.293) | | |
| $Log(PROD_j)$ | | | −1.12*** | −5.22*** |
| | | | (0.191) | (0.467) |
| $Log(DIST_{ij})$ | −0.79*** | −1.71*** | −1.16*** | −2.99*** |
| | (0.051) | (0.189) | (0.068) | (0.165) |
| $Log(WAGE_j)$ | | 1.22*** | | |
| | | (0.453) | | |
| $Log(FDIopen_j)$ | | 0.05 | | |
| | | (0.105) | | |
| $Log(KF_j)$ | | −0.83** | | |
| | | (0.422) | | |
| $Log(\frac{KBL_i}{POP_i})$ | | | | 1.61*** |
| | | | | (0.431) |
| *SIM* | −0.33*** | 1.79*** | −0.28*** | 0.03 |
| | | (0.206) | (0.073) | (0.172) |
| *RLF* | 0.01 | 0.02 | 0.04*** | −0.06*** |
| | (0.007) | (0.025) | (0.009) | (0.022) |
| *EMU* | 0.16*** | −0.75*** | −0.07** | 0.35*** |
| | (0.024) | (0.101) | (0.035) | (0.083) |
| *EAST* | −1.16*** | −3.75*** | −0.22 | 2.41*** |
| | (0.294) | (0.775) | (0.341) | (1.001) |
| *BORDER* | 0.71 | 1.04 | −1.10 | 0.90 |
| | (0.411) | (0.968) | (0.629) | (1.406) |
| *CEEC* | 0.58** | −5.53*** | −1.14*** | −6.34*** |
| | (0.293) | (0.826) | (0.393) | (1.207) |

(*continued on the next page*)

**Table 5.6** (Continued)

| Dep. variable | FEVD-SUR | | | |
|---|---|---|---|---|
| | Exports | FDI out | Imports | FDI in |
| Time effects | Yes | Yes | Yes | Yes |
| ($p$-value of Wald test) | (0.00) | (0.00) | (0.00) | (0.00) |
| No. of system observation | | | 10660 | |
| No. of obs. per equation | 2665 | 2665 | 2665 | 2665 |
| No. of groups per equation | 353 | 353 | 353 | 353 |
| $\|m\|$-stat. SUR/OLS | 9.60 | 10.39 | 63.93 | 8.92 |
| ($p$-value) | (0.97) | (0.98) | (0.00) | (0.98) |
| $\|m\|$-stat. HT-SYS/FEVD-SYS | 115.15 | 117.98 | 20.14 | 15.36 |
| ($p$-value) | (0.00) | (0.00) | (0.44) | (0.80) |
| Resid. based ADF test | 659.7** | 1418.5*** | 1185.8*** | 1027.4*** |
| ($p$-value) | (0.01) | (0.00) | (0.00) | (0.00) |
| $R^2$ | 0.53 | 0.58 | 0.63 | 0.58 |

*Note*: Standard errors are robust to heteroscedasticity, for a description of the wild bootstrap algorithm to adjust 2. step standard errors see text. The number of bootstrap repetitions is set to 1000
*Denote statistical significance at the 10% level **Denote statistical significance at the 5% level
***Denote statistical significance at the 1% level

close and explain a significant part of the total variation in the respective trade and FDI equations (around 50–70%). Taking a closer look at the individual equations' variable coefficients, we find that most key variables are estimated in line with our a-priori expectations. Output effects (both GDP for the home and foreign country) proxying the role of economic mass in bilateral trade and FDI activity play a distinct role. This is in line with our theoretical assumptions. Only for the export equation the results show a surprisingly low explanatory power of the income variables: Here the effect is mainly captured through labor productivity (defined as GDP per total employment). Econometrically, this latter result may hint at the strong link between labor productivity and export activity, which is broadly confirmed in the closely related micro-based literature (see e.g. Helpman et al. 2003; Arnold and Hussinger 2006).

All equations assign a crucial role to distance as a proxy for transportation costs in both trade/FDI, while the effect is found to be on average higher in the FDI rather than trade case. The latter result may reflect the likely path dependency in building up FDI stocks, since the rather more distant peripherical EU27 member states (from the geographical perspective of Germany) have only recently joined the EU (and thus adopted the institutional setup of the *aquis communitaire*). Moreover, the empirical result that distance exerts a stronger negative impact on foreign affiliate production than exports can be related to similar results in the recent literature (see e.g. Ekholm 1998).[21]

[21] Also Markusen and Maskus (1999) and Carr et al. (2001) among others report a significant negative influence of distance on outward FDI/foreign affiliate production.

For export activity the EMU dummy shows the a-priori expected positive impact on German exports for both estimators. From 1999 onwards, German export activity to the other EMU member states is estimated to be above its normal potential (in terms of being adjusted for economic mass, geographical distance and other explanatory variables as specified in the gravity model of (5.4)). For inward FDI, we find similar investment enhancing effects of EMU creation. The results are found to be robust for both the HT and FEVD estimator. However, on the contrary, the effect on outward FDI is found to be negative, possibly reflecting the general trend of stagnating or even decreasing German FDI stocks in the EMU countries contrary to non-EMU economies within the EU27 (especially a shift from the peripherical, southern Mediterranean EMU member states to the CEECs throughout the late 1990s). For imports, the estimated EMU coefficient turns out to be insignificant in the HT-case and only marginally negative in the FEVD-SUR approach. Also, with respect to the border dummy, we do not find any statistically significant result for both estimators.

The dummy variables for the East German states and CEEC economies turn out to be strongly negative in most specifications. For the export and outward FDI equation the East German states dummy is found to be significantly negative indicating that the macro region is still far beyond its trading potential, we would expect according to its economic mass and geographical location within the EU27.[22] On the contrary, for inward FDI equation, both estimators find a significant and positive coefficient for this dummy variable. This result mirrors the qualitative findings from the stylized facts, saying that the East German states throughout their economic transition process are limited to act as an FDI host country with little options to actively invest abroad. Moreover, the positive coefficient for the East German macro region in the inward FDI equation may reflect the large-scale investment promotion scheme for the East German economy jointly launched by the EU, federal and state level government, which significantly lowered the regional user costs of capital and led to an inflow of (foreign and West German) capital.

The results for the CEEC dummy in the export equation are somewhat mixed. While the HT model produces a (weakly significant) negative CEEC dummy, the FEVD output reports a positive coefficient sign. With respect to German exports to the CEECs, the latter positive dummy variable coefficient indicates that trade flows to these countries are above their normal potential, which has been widely confirmed in earlier empirical contributions for the first half of the 1990s.[23] On the contrary,

[22]Related to our results Alecke et al. (2003) find a significant negative dummy variable for East German states in a gravity model context for estimating German regional trade flows to Poland and Czech Republic.

[23]It remains an open question though whether this result is also expected to hold for the rapid economic catching up process of the CEECs. Moreover it is not clear whether Germany is likely to hold its first-mover advantages compared to the other EU15 countries: While Kunze and Schumacher (2003) predict a further boost in the German CEEC trade, Buch and Piazolo (2000) and Caetano et al. (2002) among others make projections based on gravity models that Germany throughout the 1990s has already exploited most of its trade potential with CEE countries, and that in the following other EU15 member states are expected to benefit most from the recent EU enlargement.

the CEEC dummy in the outward FDI equation is found to be significantly negative for both estimators indicating that German outward FDI stocks in these economies are still below their 'normal' potential. Moreover, the persistently negative CEEC dummy in the import and inward FDI equation reflect our a-priori expectations that these countries due to historical and structural reasons still have very limited capacities to export and invest abroad.

### *5.5.4 Robustness Checks*

To check for the appropriateness of our empirical specification in the HT case, we compute a weak identification test to measure the degree of instrument correlation with the endogenous regressors to identify low correlation levels, which in turn may translate into a poor overall performance (see e.g. Stock and Yogo 2005). For the HT-3SLS-GMM model, all equations pass the weak identification test in terms of the Staiger and Stock (1997) rule of thumb ($F \geq 10$). We also apply the Sargan (1958)/Hansen (1982) test for overidentification of moment conditions. The results of the overidentification test show that, except for the inward FDI model, all chosen IV sets have rather low test statistics.[24] For the inward FDI equations all attempts to further reduce the number of moment conditions above those reported in Table 5.5 result in an instability of most variable coefficients so that we rely on the reported IV set even though it fails to pass the Sargan overidentification test.

To compare the appropriateness of our chosen full information system approach relative to a limited information benchmark, we employ the Hausman (1978) test ($m$-stat.). Under the assumption that the 3SLS estimator is generally more efficient than the 2SLS estimator, we test whether the difference between the two estimators is large, indicating that the more complex GLS transformation in the 3SLS case is likely to induce a misspecification in the model rendering it inconsistent. Thus, under the null hypothesis, both estimators are consistent, but only 3SLS is efficient. Under the alternative hypothesis only 2SLS is consistent.[25] For the FEVD model we use an analogous test framework comparing the SUR approach with the OLS benchmark. The results of the Hausman test in Tables 5.5 and 5.6 show that the full information techniques (both in the HT and FEVD case) pass the test for convenient confidence intervals in all equations except for imports. In sum we take these results in favor for our specified full information techniques.

In the spirit of Baltagi et al. (2003), we also employ a second Hausman test to check for the consistency and efficiency of the HT estimator against the FEVD

[24]Since the overidentification test tends to be very restrictive in terms of hypothesis rejection, we take tests results for which the null hypothesis of instrument appropriateness is not rejected at the 1% level in favor for the respective IV set in focus.

[25]By construction, if the 2SLS variance is larger than the 3SLS variance, the test statistic will be negative. Though the original test is typically not defined for negative values, here we follow Schreiber (2007) and take the absolute value of the Hausman $m$-stat. as indicator for rejecting the null hypothesis of 3SLS efficiency.

benchmark, where the latter builds upon consistent FEM estimation for the vector of time-varying variables. We thus have a testable null hypothesis for this parameter vector, while we cannot evaluate the consistency and efficiency of the vector of time-fixed variables. The results of this second Hausman test are reported in Table 5.6 and indicate that the difference between the two estimators is rather small for the import and inward FDI equation, where the null hypothesis of consistency and efficiency of the HT model cannot be rejected for convenient confidence intervals. However, for the export and outward FDI equation the null hypothesis is clearly rejected. Taken together with the empirical findings in Mitze (2009) that Hausman–Taylor type models tend to have a severe bias in estimating the coefficient vector of time-fixed variables, we favor the FEVD-SUR approach for our empirical application since it less sensitive to likely problems in IV selection. Finally, as indicated by the residual based ADF-test for cointegration in the spirit of Kao (1999), for both models we can reject the null hypothesis for non-stationarity in the residuals.

## 5.6 Identification of Trade-FDI Linkages

We find significant cross-equation correlations for both estimators. Given the favoring postestimation results from above we favor the FEVD-SUR estimates, which are nevertheless qualitatively broadly in line with the Hausman–Taylor results.[26] In Table 5.7 we plot the corresponding (rank) correlation coefficients for our four-equation residual based VCV matrix together with the Breusch–Pagan LM test results for unbalanced data. Additionally, we also compute a Harvey–Phillips (1982) type exact independence $F$-test, which checks for the joint significance of the other equations' residuals in an augmented first step regression (see e.g. Dufour and Khalaf 2002, for details).

We get significant evidence for both substitutive and complementary linkages among the variables under observation. Focusing on each type of international activity separately, for both the exports and imports as well as outward and inward FDI activity respectively we observe complementary (enhancing) effects. Turning to the trade-FDI linkages we find a substitutive relationship between exports and outward FDI activity in line with earlier evidence reported in Jungmittag (1995) as well as Egger and Pfaffermayr (2004). Also, imports and outward FDI are found to be of substitutive nature. However, on the contrary imports and inward FDI are found to complement each other, while the relationship between exports and inward FDI is tested insignificantly on the basis of Breusch–Pagan LM tests. As a sensitivity analysis we also estimate trade-FDI linkages for sub-aggregates of our data set as:

- West Germany—EU27/EU15,
- East Germany—EU27/EU15.[27]

[26] Results for the latter estimator can be obtained upon request from the authors.

[27] A further disaggregation is not feasible due to data limitations.

**Table 5.7** Cross-equation residual correlation and Breusch–Pagan test for German—EU27

| | Exports | FDI out | Imports | FDI in |
|---|---|---|---|---|
| Exports | 1.00 | | | |
| FDI out | $-0.44^{***}$ | 1.00 | | |
| | $\chi^2(1) = 71.9$ | | | |
| Imports | $0.53^{***}$ | $-0.15^{***}$ | 1.00 | |
| | $\chi^2(1) = 95.5$ | $\chi^2(1) = 8.69$ | | |
| FDI in | 0.02 | $0.25^{***}$ | $0.41^{***}$ | 1.00 |
| | $\chi^2(1) = 0.12$ | $\chi^2(1) = 27.3$ | $\chi^2(1) = 62.1$ | |
| Harvey–Phillips ($p$-val.) | (0.00) | (0.00) | (0.00) | (0.00) |

*Denote statistical significance at the 10% level **Denote statistical significance at the 5% level ***Denote statistical significance at the 1% level

**Table 5.8** Cross-equation residual correlation and Breusch–Pagan test for West German—EU27

| | Exports | FDI out | Imports | FDI in |
|---|---|---|---|---|
| Exports | 1.00 | | | |
| FDI out | $-0.16^{**}$ | 1.00 | | |
| | $\chi^2(1) = 4.01$ | | | |
| Imports | $0.33^{***}$ | $0.19^{***}$ | 1.00 | |
| | $\chi^2(1) = 43.8$ | $\chi^2(1) = 24.2$ | | |
| FDI in | $0.14^{***}$ | $0.35^{***}$ | $0.71^{***}$ | 1.00 |
| | $\chi^2(1) = 9.69$ | $\chi^2(1) = 53.7$ | $\chi^2(1) = 140.9$ | |
| Harvey–Phillips ($p$-val.) | (0.00) | (0.00) | (0.00) | (0.00) |

*Denote statistical significance at the 10% level **Denote statistical significance at the 5% level ***Denote statistical significance at the 1% level

Our motivation for using these additional subsamples is that the data period from 1993–2005 covers the transformation period of the central and eastern European countries (including also the East German economy) from planned to market economies. Given the historical situation of these countries, we only observe a gradual opening up for internationalization activities with the core EU-15 member states over the sample period, which may well impact on the empirical results. We thus expect that trade-FDI ties are supposed to be strongest for the West German states with their respective EU-15 bilateral country pairs.

In Table 5.8, we see that the identified cross-equation correlations closely follow predictions of New Trade theory models such as Baldwin and Ottaviano (2001). That is, when international trade is merely of intra-industry type with non-zero trade costs, the latter shifts production abroad and leads to export replacement effects of FDI. However, at the same time FDI may stimulate trade via reverse good imports.

**Table 5.9** Cross-equation residual correlation and Breusch–Pagan test for West German—EU15

| | Exports | FDI out | Imports | FDI in |
|---|---|---|---|---|
| Exports | 1.00 | | | |
| FDI out | 0.30*** | 1.00 | | |
| | $\chi^2(1) = 49.7$ | | | |
| Imports | 0.66*** | 0.13*** | 1.00 | |
| | $\chi^2(1) = 124.5$ | $\chi^2(1) = 9.67$ | | |
| FDI in | 0.10*** | 0.75*** | −0.03 | 1.00 |
| | $\chi^2(1) = 7.80$ | $\chi^2(1) = 150.7$ | $\chi^2(1) = 0.33$ | |
| Harvey–Phillips ($p$-val.) | (0.00) | (0.00) | (0.00) | (0.00) |

*Denote statistical significance at the 10% level **Denote statistical significance at the 5% level
***Denote statistical significance at the 1% level

We thus find that export and outward FDI activity are still substitutes. However, all remaining trade-FDI links show complementary effects. In the model of Baldwin and Ottaviano (2001), this result is mainly driven by cross-hauling of FDI generating reciprocal trade effects in differentiated final products. Given the dominance of intra industry trade and horizontal FDI between West Germany and the EU27 economies as well as non-zero trade costs (as tested in our gravity model), these theoretical predictions may be seen as a good explanation for our empirically identified trade-FDI nexus in the case of West Germany. Moreover, a further disaggregation to West German—EU15 trade and FDI activity in Table 5.9 even reveals complementaries among export and FDI activity, which have not been identified for German data before, but generally match the mainstream empirical evidence in an international perspective. The latter result may be explained by the greater similarities in levels of development of West Germany and the EU15 compared to the enlarged EU including the new eastern member states, which is likely to have an effect on the horizontal/vertical nature of FDI. For the results for the East German macro region in Tables 5.10 and 5.11, we find merely substitutive linkages (except for inward FDI and trade in the East German—EU15 case), which may hint at the rather low level of internationalization activities (in particular outward FDI) of the East German macro region. Moreover, as for the West also for East Germany selective structural differences between the EU15 and the EU27 samples can be observed (e.g. with respect to inward FDI and trade variables), which may indicate the specific relation of East Germany with respect to the new Eastern EU member states.

To sum up, in addition to recent findings supporting the need of a sectoral disaggregation in analyzing trade-FDI linkages (e.g. Pfaffermayr 1996; Bloningen 2001; Türkcan 2007), our results show that the regional perspective within a nation's trade and FDI activity may also be of great importance in identifying cross-variable linkages. That is, while we find that the relationship between exports and inward FDI is found to insignificant at the aggregate level, regionally we find opposing effects

**Table 5.10** Cross-equation residual correlation and Breusch–Pagan test for East German—EU27

| | Exports | FDI out | Imports | FDI in |
|---|---|---|---|---|
| Exports | 1.00 | | | |
| FDI out | −0.48*** | 1.00 | | |
| | $\chi^2(1) = 67.6$ | | | |
| Imports | 0.80*** | −0.44*** | 1.00 | |
| | $\chi^2(1) = 161.2$ | $\chi^2(1) = 58.4$ | | |
| FDI in | −0.56*** | 0.35*** | −0.55*** | 1.00 |
| | $\chi^2(1) = 113.8$ | $\chi^2(1) = 44.1$ | $\chi^2(1) = 113.7$ | |
| Harvey–Phillips ($p$-val.) | (0.00) | (0.00) | (0.00) | (0.00) |

*Denote statistical significance at the 10% level **Denote statistical significance at the 5% level
***Denote statistical significance at the 1% level

**Table 5.11** Cross-equation residual correlation and Breusch–Pagan test for East German—EU15

| | Exports | FDI out | Imports | FDI in |
|---|---|---|---|---|
| Exports | 1.00 | | | |
| FDI out | −0.44*** | 1.00 | | |
| | $\chi^2(1) = 75.5$ | | | |
| Imports | 0.77*** | −0.45*** | 1.00 | |
| | $\chi^2(1) = 168.9$ | $\chi^2(1) = 74.6$ | | |
| FDI in | 0.76*** | −0.40*** | 0.69*** | 1.00 |
| | $\chi^2(1) = 161.6$ | $\chi^2(1) = 62.3$ | $\chi^2(1) = 152.9$ | |
| Harvey–Phillips ($p$-val.) | (0.00) | (0.00) | (0.00) | (0.00) |

*Denote statistical significance at the 10% level **Denote statistical significance at the 5% level
***Denote statistical significance at the 1% level

(a positive one between West Germany—EU27, a negative one for East Germany—EU27) which on average may cancel out a total net effect. A similar interpretation can be given to the strong negative correlation between exports and outward FDI in the case of East Germany, which is likely to influence the aggregate results. This latter result may especially stem from the fact that for our sample period, the dynamics of integration to world markets for East Germany is much higher due to its low starting levels and putting distinct choice option on the mode of internationalization.[28] The identified trade-FDI linkages are shown in Table 5.12.

[28]It is not clear whether this result can be captured in a level effect, or whether the assumption of slope homogeneity for the time varying variables is not valid for the underlying German regions (see e.g. Pesaran and Yamagata (2008)). Future research should put more effort on this question, especially when longer time dimensions of the variables are available.

**Table 5.12** Identified trade-FDI linkages for different data samples

| | Exports | FDI out | Imports | FDI in |
|---|---|---|---|---|
| Germany—EU27 | | | | |
| Exports | * | | | |
| FDI out | negative | * | | |
| Imports | positive | negative | * | |
| FDI in | insign. | positive | positive | * |
| West Germany—EU27 | | | | |
| Exports | * | | | |
| FDI out | negative | * | | |
| Imports | positive | positive | * | |
| FDI in | positive | positive | positive | * |
| West Germany—EU15 | | | | |
| Exports | * | | | |
| FDI out | positive | * | | |
| Imports | positive | positive | * | |
| FDI in | positive | positive | insign. | * |
| East Germany—EU27 | | | | |
| Exports | * | | | |
| FDI out | negative | * | | |
| Imports | positive | negative | * | |
| FDI in | negative | positive | negative | * |
| East Germany—EU15 | | | | |
| Exports | * | | | |
| FDI out | negative | * | | |
| Imports | positive | negative | * | |
| FDI in | positive | negative | positive | * |

## 5.7 Conclusion

The aim of this chapter was to analyze the main macroeconomic driving forces for German regional and national trade and FDI activity within the EU27 and to identify their correlations. We have used the gravity approach as a modelling framework and base our identification strategy on the inclusion of appropriate exogenous control variables as proposed in the gravity model literature. With respect to the underlying trade-FDI linkages at the aggregate level, we basically find a substitutive relationship between exports and outward FDI activity in line with earlier evidence reported in Jungmittag (1995) as well as Egger and Pfaffermayr (2004). Also, imports and outward FDI are found to be substitutive, while imports and inward FDI complement each other.

We also estimated trade-FDI links for regional sub-samples. That is, for West German—EU27 trade/FDI activity, we find strong support for the predictions of NTT models as in Baldwin and Ottaviano (2001). When international trade is of merely intra-industry type with non-zero trade costs, the latter shifts production abroad and leads to export replacement effects of FDI. However, at the same time FDI may stimulate trade via reverse good imports. Thus, export and outward FDI are found to be substitutes for each other, while all remaining variable linkages show complementary effects. The latter result may indicate the growing importance of vertical FDI in our sample period from 1993 to 2005, which may be especially driven by a boost of investment activity in the new EU member states. Moreover, a further disaggregation into West German—EU15 trade/FDI activity even reveals complementaries among export and FDI activity, which have not been identified for German data before, but match with the general empirical evidence in an international context. For the East German states, we overwhelmingly find substitutive linkages (except for inward FDI and trade in the East German—EU15 case), which may indicate the rather low level of internationalization activities (in particular outward FDI) of the East German macro region.

When interpreting these results, we have to account for our chosen country sample and time period. While our results make sense for intra-EU trade and FDI activity, a generalization to overall trade-FDI activity has to be done carefully.[29] These caveats have to be taken into account when the results are used in the policy debate for export and/or FDI promotion schemes. Our results also indicate to look at regional disaggregation when modelling trade and FDI patterns and identifying underlying cross-variable linkages. Future research effort should be done in explicitly testing for the significance of other factors driving internationalization activity besides those already captured in our approach (such as exchange rates) as well as to more carefully account for the likely caveats when operationalizing the gravity model. This latter point may comprise explicit tests for the poolability of the data (see e.g. Pesaran and Yamagata 2008) as well as the appropriate functional form.

## Appendix A: IV and Non-IV System Estimators

### *A.1 The General Model*

We start from a general, triple indexed model form as:

$$y_{ijt} = \alpha + \beta' X_{ijt} + \gamma' Z_{ij} + u_{ijt} \quad \text{with } u_{ijt} = \mu_{ij} + \nu_{ijt}, \tag{5.8}$$

with $i = 1, 2, \dots, N$; $j = 1, 2, \dots, M$ and $t = 1, 2, \dots, T$. The endogenous variable ($y_{ijt}$) and the vector of time varying explanatory variables ($X_{ijt}$) may vary in all

[29] Even though German-EU27 trade and FDI pattern accounts for a large share of total trade and FDI activity. Moreover, using a world sample Cechella et al. (2008) recently found that world FDI is also mainly driven by horizontal motives.

three dimensions of our model, while the vector of time fixed explanatory variables $(Z_{ij})$ is kept constant across $t$. $\beta$ and $\gamma$ are vectors of regression coefficients, $\alpha$ is the overall constant term and $u_{ijt}$ is the composed error term including the unobservable individual effects $\mu_{ij}$ and a remainder error term $\nu_{ijt}$. Typically the latter two are assumed to be i.i.d. residuals with zero mean and constant variance. For system estimation we may write (5.8) compactly as:

$$y_n = R_n \xi_n + u_n, \quad u_n = \mu_n + \nu_n, \tag{5.9}$$

where $n$ denotes the $n$th structural equation of the system with $n = 1, \ldots, M$. In our case $M = 4$. $R_n = (X_n, Z_n)$ and $\xi = (\beta', \gamma')$. Following Cornwell et al. (1992) we then simply stack the equations into the usual 'starred' form as:

$$y_* = R_* \xi_* + u_*, \tag{5.10}$$

where $y'_* = (y'_1, \ldots, y'_N)$ and similar for $\xi_*$ and $u_*$. $R_*$ is defined as

$$R_* = \begin{bmatrix} R_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & R_M \end{bmatrix}. \tag{5.11}$$

Depending on the type of estimator we can make use of the seemingly unrelated regression (SUR) approach or 3SLS estimation to the stacked system in (5.10). Thereby, the SUR model may be seen as a special case of the more general 3SLS estimator when there is no right hand side endogeneity in the estimated equations (for details see e.g. Intrilligator et al. 1996). The SUR approach is popular since it captures the correlation of the disturbances across equations and—if the disturbance terms are correlated—it is asymptotically more efficient than OLS for each single equation. However, for the case we have to cope with endogeneity of the right-hand side regressors of the model either in the sense of endogenous variables as explanatory variables in other equations of the system or a correlation of some regressors with the disturbances, Baltagi (2008) proposes to use 3SLS for estimating (5.10).

## *A.2 The HT-3SLS-GMM Estimator*

Since the logic of the Hausman–Taylor model centers around consistent IV estimation of all parameters in the model, the 3SLS estimator is the natural choice (or in a broader context system GMM).[30] Next to consistent IV choice for estimation purposes one also has to decide about the proper empirical form of the system's error term variance-covariance matrix. In its standard form the model typically builds on the random effects assumption in line with Baltagi's (1981) feasible EC-3SLS es-

[30] The system extension to the standard single equation Hausman–Taylor models was first proposed by Cornwell et al. (1992), a GMM version of the estimator is discussed in Ahn and Schmidt (1999).

timators as probably the most prominent example in the field of system estimation with Panel data. As Cornwell et al. (1992) show, the EC-3SLS estimator can be interpreted as a special form of the more general HT-3SLS framework, namely when all exogenous variables are assumed to be independent of the system's error components. Alternatively, Ahn and Schmidt (1999) propose to start with an unrestricted covariance matrix in the context of optimal system GMM estimation and then test for valid model (variance-covariance) restrictions. For the purpose of this analysis we specify the Hausman–Taylor model in its 3SLS-GMM form as:

$$\hat{\beta}_{3SLS\text{-}GMM} = [R_*' H_* (H_*' \hat{\Omega} H_*)^{-1} H_*' R_*]^{-1} R_*' H_* (H_*' \hat{\Omega} H_*)^{-1} H_*' y_*, \quad (5.12)$$

where $H_*^S$ is the system's total IV set based on the definition $H_i^S = I_M \otimes H_i$ (with $H_i$ as the $n$th equation instrument set) and $u_i^S = (u_{1i}', \ldots, u_{M,i}')$, so that we can write the system's overall set of moment conditions compactly as $E(H_i^{S'} u_i^S) = 0$. The latter in turn is chosen according to th Hausman–Taylor assumptions. $\hat{\Omega} = \text{Cov}(u_*)$ is the variance-covariance matrix of the equation system. The main difference between the standard 3SLS estimator and its 3SLS-GMM alternative is that the latter allows for different instruments in subsequent equations, while standard 3SLS estimation assumes the same IV-set applies to every equation in the system. The latter assumption may be somewhat problematic in our case, since we have found that different instruments are valid for subsequent model equations based on a series of Hansen (1982)/Sargan (1958) overidentification tests for the single equation benchmark models.[31]

For convenience and in line with the mainstream literature on the Hausman–Taylor model we assume that $\Omega_*$ takes the random effect form.[32] We thus model the two error components $\mu$ and $\nu$ as i.i.d. with $(0, \Sigma_\mu)$ and $(0, \Sigma_\nu)$, where $\Sigma_\mu = [\sigma^2_{\mu(j,l)}]$ is the $4 \times 4$ variance-covariance matrix corresponding to the unobserved individual effects (with $j, l =$ [exports, FDI out, imports, FDI in]) and $\Sigma_\nu = [\sigma^2_{\nu(j,l)}]$ is the $4 \times 4$ variance-covariance matrix of the remainder error term. For unbalanced panel data the variance-covariance varies with $ij$ and therefore transforming the estimation system by $\Omega_{ij}^{-1/2}$ takes the following form:

$$\Omega_{ij}^{-1/2} = (\Sigma_\nu + T_{ij} \Sigma_\mu)^{-1/2} \otimes P + \Sigma_\nu^{-1/2} \otimes Q. \quad (5.13)$$

In empirical terms we use the feasible GLS approximation in order to replace the unknown parameters of covariance matrix, $\Sigma_\nu$ and $(\Sigma_\nu + T_{ij} \Sigma_\mu)$ by consistent estimates. To derive these proxies we follow Baltagi's (2008) suggestion for unbalanced panels and estimate the respective sub blocks (or matrix elements) of $\hat{\Sigma}_\nu$ and $\hat{\Sigma}_\mu$ as

[31] Results can be obtained upon request from the authors.

[32] An alternative choice for $\Omega_*$ would be an unrestricted form in analogy to the optimal weighting matrix for system GMM as $\Omega = (I_N \otimes \Sigma_{j,l})$, where $\Sigma_{j,l}$ can be estimated from any consistent 1.step residuals according to $\Sigma_{j,l} = N^{-1} \sum_{i=1,j=1}^{NM} (\hat{u}_j \hat{u}_l')$ (see Ahn and Schmidt 1999, for details).

$$\hat{\sigma}^2_{\nu_{(j,l)}} = \frac{\hat{u}'_{j,l} Q \hat{u}_{j,l}}{\sum_{i=1,j=1}^{NM}(T_{ij}-1)}, \tag{5.14}$$

$$\hat{\sigma}^2_{\mu_{(j,l)}} = \frac{\hat{u}'_{j,l} P \hat{u}_{j,l} - NM\hat{\sigma}_{\nu_{(j,l)}}}{\sum_{i=1,j=1}^{NM}(T_{ij})}, \tag{5.15}$$

where $\hat{u}$ is the estimation residual from an untransformed 1. step 2SLS estimation (see also Baltagi 2008, or Baltagi and Chang 2000, for details).[33]

## A.3 *The FEVD(-SUR) Estimator and Bootstrapping Standard Errors*

An alternative to the Hausman–Taylor IV-estimator is an augmented FEM approach proposed by Plümper and Tröger (2007) for the single equation case. The goal of the so-called Fixed Effects Vector Decomposition (FEVD) model is to run a consistent FEM model and still get estimates for the time-invariant variables. The intuition behind FEVD specification is as follows: The unobservable individual effects are a vector of the mean effect of omitted variables, including the effect of time-invariant variables. According to Plümper and Tröger (2007) it is therefore possible to regress the proxy for individual effects derived from the FEM residuals on the time-invariant variables to obtain approximate estimates for these variables. The estimator builds on the following steps: First, we apply a standard FEM on (5.8) to obtain the vector of time-varying variable $\beta$. Second, we use the estimated vector of group residuals as proxy for the unobservable individual effects $\hat{\mu}_{ij}$ to run a regression of the explanatory time-fixed variables against this 'generated regressand' as:

$$\hat{\mu}_{ij} = \omega + \hat{\delta}' Z_{ij} + \eta_{ij}, \tag{5.16}$$

where $\omega$ is an overall intercept and $\eta_{ij}$ is the residual. The second step aims at identifying the unobserved parts of the individual effects. In a third (optional) step Plümper and Tröger re-estimate (5.8) in a POLS setup including the 2. step residual $\eta_{ij}$ to control for collinearity between time-varying and time-fixed right hand side variables. Finally, it is important that standard errors for the time-fixed variable coefficients have to be corrected due to the use of a 'generated regressand' in the 2. modelling step to avoid an overestimation of $t$-values. To sum up, the FEVD 'decomposes' the estimated proxy for the unobservable individual effects obtained from the FEM residuals into one part explained by the time-fixed variables and a remainder error term. Plümper and Tröger argue that one major advantage of the FEVD compared to the Hausman–Taylor model is that there is no need for any arbitrary ex-ante variable classification for consistent IV selection.

[33] Finally, in the system transformation process we follow Baltagi (2008) and apply the Cholesky decomposition to $\Sigma_{\nu}^{-1}$ and $\Sigma_{\mu}^{-1}$.

However, as shown in Mitze (2009) although the researcher is not confronted with the choice of classifying variables as being exogenous or endogenous with respect to the error term, the FEVD itself makes an implicit choice: That is, in specifying the time-varying variables the model follows the generality of the FEM approach, which assumes a variable correlation of unknown form. With respect to the time invariant variables the estimator on the other hand assumes in its basic form that none of the time-fixed variable is correlated with the individual effects.[34] If the implicit (and fixed) choice of the FEVD does not reflect the true correlation between the variables and the error term the estimator may perform poor. However, Monte Carlo simulations by Alfaro (2006), Plümper and Tröger (2007) and Mitze (2009) show that even if the FEVD does not meet the underlying true orthogonality conditions of the data set, due to is robust non-IV specification it has a smaller bias and prediction errors than consistent Hausman–Taylor specification especially for estimating the coefficients of both endogenous and exogenous time-fixed variables.

As outlined in Sect. 5.4, the system extension to the FEVD is rather straightforward. To correct standard errors in the resulting FEVD-SUR approach we apply the 'wild bootstrap' technique, which is implemented through the following steps as outlined in Atkinson and Cornwell (2006):[35]

**Step 1** Estimate the coefficient vector $\hat{\beta}_{FEM\text{-}SUR}$ of $X_{it}$ in a SUR system based on the within-type transformed data (FEM).

**Step 2** Using the coefficient vector $\hat{\beta}_{FEM\text{-}SUR}$, we compute

$$\hat{\pi}_i = \bar{y} - \hat{\beta}_{FEM\text{-}SUR}\bar{X}_i. \tag{5.17}$$

**Step 3** Estimate the coefficient vector $\hat{\gamma}_{POLS\text{-}SUR}$ for $Z_i$ by POLS-SUR.

**Step 4** Compute the second step residuals as

$$\hat{\xi}_{it} = y_{it} - \hat{\beta}_{FEM\text{-}SUR}X_{it} - \hat{\gamma}_{POLS\text{-}SUR}(J_T \otimes Z_i). \tag{5.18}$$

According to the 'wild bootstrap' procedure replace $\hat{\xi}_{it}$ with

$$\tilde{\xi}_{it} = (\hat{\xi}_{it})\tilde{\upsilon}_{it} \quad \text{where } f(\hat{\xi}_{it}) = \frac{\hat{\xi}_{it}}{(1-h_{it})^{1/2}} \tag{5.19}$$

and $h$ is the model's projection matrix so that a division by $(1-h_{it})^{1/2}$ ensures that the transformed residuals have the same variance (for details see MacKinnon 2002); $\tilde{\upsilon}_{it}$ is defined as a two-point distribution (the so-called Rademacher distribution) with

$$\tilde{\upsilon}_{it} = \begin{cases} -1 & \text{with probability } 1/2, \\ \phantom{-}1 & \text{with probability } 1/2. \end{cases} \tag{5.20}$$

[34] In fact, a modification of the FEVD also allows for the possibility to estimate the second step as IV regression and thus account for endogeneity among time invariant variables and $\eta_{ij}$. However, this brings back the classification problem from the Hausman–Taylor specification, which we explicitly aim to avoid by non-IV estimation.

[35] For notational convenience the cross-section dimension is expressed by $i$ rather than $ij$ here.

**Step 5** For each of $i = 1, \ldots, N$ blocks, we draw randomly with replacement $T$ observations with probability $1/T$ from $\tilde{\upsilon}_{it}$ to obtain $\tilde{\xi}^*_{it}$.
**Step 6** Generate

$$y^*_{it} = \hat{\beta}_{FEM\text{-}SUR} X_{it} - \hat{\gamma}_{POLS\text{-}SUR}(J_T \otimes Z_i) + \tilde{\xi}^*_{it}. \tag{5.21}$$

**Step 7** Compute the FEM-SUR for the vector of variable coefficients $\beta$ using the starred data as $\beta^*_{FEM\text{-}SUR}$.
**Step 8** Using $\beta^*_{FEM\text{-}SUR}$ from the previous step to compute

$$\omega_i = \bar{\tilde{\xi}}_i - (\hat{\beta}^*_{FEM\text{-}SUR} - \hat{\beta}_{FEM\text{-}SUR})\bar{X}_i. \tag{5.22}$$

**Step 9** Randomly resample with replacement from $\hat{u}_i$ to obtain $u^*_i$. Then compute

$$\pi^*_i = \hat{\gamma}_{POLS\text{-}SUR} Z_i + u^*_i. \tag{5.23}$$

**Step 10** Estimate the coefficients $\gamma^*_{POLS\text{-}SUR}$ using the starred data.
**Step 11** Repeat steps 5–9 1000 times and compute the sample standard deviation of $\gamma^*_{POLS\text{-}SUR}$ as an estimator of the standard error of $\hat{\gamma}_{POLS\text{-}SUR}$.

## Appendix B: Testing for Cross-Equation Residual Correlation

In order to analyze the statistical significance of the identified cross-equation residual correlation we use Breusch–Pagan (1980) type tests corrected for unbalanced panel data sets according to Song and Jung (2001) and Baltagi and Song (2006).[36] The Breusch–Pagan LM test on the correlation of individual effects across equations can be defined as

$$BP = \left(\frac{1}{2}\right) n^2 [A^2/(J - n)], \tag{5.24}$$

$$\text{with } J = \sum_{i=1, j=1}^{NM} T_{ij} \times (T_{ij} - 1),$$
$$A = [(u_j \Delta_1 \Delta'_1 u_l)/((u'_j u_j)(u'_l u_l))^{1/2}],$$
$$\Delta_1 = (D'_1, D'_2, \ldots, D'_T)',$$

where $n$ is the number of total observations and $D_t$ is obtained from an identity matrix $I_{NM}$ by omitting the rows corresponding to individuals not observed in year $t$ (with $j, l =$ [exports, FDI out, imports, FDI in]). As Baltagi (2008) shows, this can be easily done by restacking the residuals such that all the individuals observed in the first period are stacked on top of those observed in the second period, and so on. In this case, the slower index is $t$ and the faster index is $i$, the error term (in vector form) can be written as $u = \Delta_1 \mu + \nu$. Testing for the cross-equation correlation

[36] Rather than using one-sided Honda (1985) type tests as proposed by Egger and Pfaffermayr (2004), since the cross equation covariance elements can actually become negative.

of the overall error term, $\Delta_1\Delta_1'$ chancels out (see e.g. Dufour and Khalaf 2002). Under the null hypothesis of no correlation, the Breusch–Pagan type LM test given by (5.24) is asymptotically distributed as $\chi^2(1)$.

## References


Africano, A., & Magalhaes, M. (2005). *FDI and trade in Portugal: a gravity analysis* (FEP Working Papers 174). University of Porto.

Aizenman, J., & Noy, I. (2006). FDI and trade-two-way linkages? *The Quarterly Review of Economics and Finance*, *46*(3), 317–337.

Ahn, S., & Schmidt, P. (1999). Estimation of linear panel data models using GMM. In L. Matyas (Ed.), *Generalized methods of moments estimation*. Cambridge.

Alecke, B., Mitze, T., & Untiedt, G. (2003). Das Handelsvolumen der ostdeutschen Bundesländer mit Polen und Tschechien im Zuge der EU-Osterweiterung: Ergebnisse auf Basis eines Gravitationsmodells. *DIW Vierteljahrshefte zur Wirtschaftsforschung*, *72*(4), 565–578.

Alfaro, R. (2006). *Application of the symmetrically normalized IV estimator* (Working Paper). Boston University.

Anderson, P. (1979). A theoretical foundation for the gravity equation. *American Economic Review*, *69*, 106–116.

Anderson, J., & van Wincoop, E. (2003). Gravity with gravitas: a solution to the border problem. *American Economic Review*, *93*(1), 170–192.

Arnold, J., & Hussinger, K. (2006). *Export versus FDI in German manufacturing: firm performance and participation in international markets* (Deutsche Bundesbank Discussion Paper No. 04/2006).

Atkinson, S., & Cornwell, C. (2006). *Inference in two-step panel data models with time-invariant regressors* (Working Paper). University of Georgia.

Baldwin, R. (1994). *Towards an integrated Europe*. London: CEPR.

Baldwin, R., & Ottaviano, G. (2001). Multiproduct multinationals and reciprocal FDI dumping. *Journal of International Economics*, *55*, 429–448.

Baldwin, R., & Taglioni, D. (2006). *Gravity for dummies and dummies for gravity equations* (NBER Working Paper No. 12516).

Baltagi, B. (1981). Simultaneous equations with error components. *Journal of Econometrics*, *17*, 189–200.

Baltagi, B. (2008). *Econometric analysis of panel data* (4th ed.). Chichester: John Wiley & Sons.

Baltagi, B., Bresson, G., & Pirotte, A. (2003). Fixed effects, random effects or Hausman Taylor? A pretest estimator. *Economic Letters*, *79*, 361–369.

Baltagi, B., & Chang, Y. (2000). Simultaneous equations with incomplete panels. *Econometric Theory*, *16*, 269–279.

Baltagi, B., & Song, S. (2006). Unbalanced panel data: a survey. *Statistical Papers*, *47*, 493–523.

Bayoumi, T., & Lipworth, G. (1997). *Japanese foreign direct investments and regional trade* (IMF Working Paper No. WP/97/103).

Belke, A., & Spies, J. (2008). Enlarging the EMU to the East: what effects on trade? *Empirica*, *4*, 369–389.

Bergstrand, J. (1985). The gravity equation in international trade: some microeconomic foundations and empirical evidence. *The Review of Economics and Statistics*, *67*(3), 474–481.

Bergstrand, J. (1989). The generalized gravity equation, monopolistic competition, and factor proportions theory of international trade. *Review of Economics and Statistics*, *71*, 143–153.

Blanchard, P., Gaigne, C., & Mathieu, C. (2008). Foreign direct investment: lessons from panel data. In L. Matyas & P. Sevestre (Eds.), *The econometrics of panel data. Fundamentals and recent developments in theory and practice* (3rd ed.). Berlin: Springer.

Bloningen, B. (2001). In search of substitution between foreign production and exports. *Journal of International Economics*, *53*, 81–104.

Borrmann, C., Jungnickel, R., & Keller, D. (2005). *What gravity models can tell us about the position of German FDI in Central and Eastern Europe* (HWWA Discussion Paper No. 328).

Brainard, S. L. (1993). *A simple theory of multinational corporations and trade with a trade-off between proximity and concentration* (NBER Working Paper No. 4269).

Brainard, S. L. (1997). An empirical assessment of the proximity-concentration tradeoff between multinational sales and trade. *American Economic Review*, *87*(4), 520–544.

Brenton, P., Di Mauro, F., & Luecke, M. (1999). Economic integration and FDI: an empirical analysis of foreign investment in the EU and in Central and Eastern Europe. *Empirica*, *26*(2), 95–121.

Breusch, T., & Pagan, A. (1980). The Lagrange multiplier test and its applications to model specification in econometrics, *Review of Economic Studies*, *47*, 239–253.

Buch, C., & Piazolo, D. (2000). *Capital and trade flows in Europe and the impact of enlargement* (Kiel Working Papers No. 1001). Kiel Institute for the World Economy.

Caetano, J., & Gallego, A. (2003). *An analysis of actual and potential trade between the EU countries and the Eastern European countries* (Documento de trabalho No. 03/2003). Universidade de Évora.

Caetano, J., Galego, A., Vaz, E., Vieira, C., & Vieira, I. (2002). *The Eastern enlargement of the Eurozone. Trade and FDI* (Ezoneplus Working Paper No. 7).

Caporale, G., Rault, C., Sova, R., & Sova, A. (2008). *On the bilateral trade effects of free trade agreements between the EU-15 and the CEEC-4 Countries* (CESifo Working Paper Series No. 2419).

Carr, D., Markusen, J., & Maskus, K. (2001). Estimating the knowledge-capital model of the multinational enterprise. *American Economic Review*, *91*, 693–708.

Cechella, C., Ramos Silva, J., & Ponce Dentinho, T. (2008). *Explaining trade and FDI relationship in a gravitational model*. Paper presented at RSAI world congress, Sao Paulo.

Cheng, I., & Wall, H. (2002). *Controlling heterogeneity in gravity models of trade* (Working Paper 1999-010C). Federal Reserve Bank of St. Louis.

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance*, *20*(2), 249–272.

Clausing, K. (2000). Does multinational activity displaces trade? *Economic Inquiry*, *38*, 190–205.

Coe, D., & Tamirisa, N. (2002). *The missing globalization puzzle* (IMF Working Paper No. WP/02/171).

Collins, S., & Rodrik, D. (1991). *Eastern Europe and the Soviet Union in the world economy*. Institute for International Economics.

Cornwell, C., Schmidt, P., & Wyhowski, D. (1992). Simultaneous equations and panel data. *Journal of Econometrics*, *51*, 151–181.

Davidson, R., & Flachaire (2001). *The wild bootstrap, tamed at last* (GREQAM Working Paper No. 99A32).

Deardorff, A. (1998). Determinants of bilateral trade: does gravity work in a neoclassical world? In J. Frenkel (Ed.), *The regionalization of the world economy* (pp. 7–33). Chicago.

Destatis (2008). *Außenhandel nach Bundesländern*, various issues. German Statistical Office, Wiesbaden.

Deutsche Bundesbank (2008). *Deutsche Direktinvestitionen im Ausland und ausländische Direktinvestitionen in Deutschland nach Bundesländern*, Statistical Publication 10, Bestandserhebungen über Direktinvestitionen S130, various issues. Frankfurt.

Dufour, J.-M., & Khalaf, L. (2002). Exact tests for contemporaneous correlation of disturbances in seemingly unrelated regressions. *Journal of Econometrics*, *106*, 143–170.

Egger, P. (2000). A note on the proper econometric specification of the gravity equation. *Economics Letters*, *66*, 25–31.

Egger, P. (2001). *SUR estimation of error components models with AR(1) disturbances and unobserved endogenous effects* (WIFO Working Paper No. 171).

Egger, P., & Pfaffermayr, M. (2004). Distance, trade and FDI: a Hausman–Taylor SUR approach. *Journal of Applied Econometrics*, *19*, 227–246.

Ekholm, K. (1998). Proximity advantages, scale economies, and the location of production. In P. Braunerhjelm & K. Ekholm (Eds.), *The geography of multinational firms* (pp. 59–76). Dordrecht: Springer.

EU Commission (2008). *AMECO database*, directorate general for economic and financial affairs, available at: http://ec.europa.eu/economy_finance/ameco.

Eurostat (2008). *National accounts (including GDP)*, various issues, available at: http://epp.eurostat.ec.europa.eu.

Feenstra, R. (2004). *Advanced international trade. Theory and evidence*. Princeton: Princeton University Press.

Fidrmuc, J. (2009). Gravity models in integrated panels. *Empirical Economics*, *37*, 435–446.

Fontagne, L., & Pajot, M. (1997). *How foreign direct investment affects international trade and competitiveness: an empirical assessment* (CEPII Working Paper No. 17).

GGDC (2008). *Total economy database, Groningen growth and development centre*, available at: http://www.ggdc.net.

Goldberg, L., & Klein, M. (1999). *International trade and factor mobility: an empirical investigation* (NBER Working Paper No. 7196).

Graham, E., & Liu, L. G. (1998). *The relationship between trade and foreign investment: empirical results for Taiwan and South Korea* (Working Paper 98-7). Institute of International Economics, Washington DC.

Graham, E. (1999). *On the relationship among direct investment and international trade in the manufacturing sector: empirical results from the U.S. and Japan* (Working Paper). Institute of International Economics, Washington DC.

Hamilton, C., & Winters, C. (1992). Opening up international trade with Eastern Europe. *Economic Policy*, *14*, 77–116.

Hansen, L. (1982). Large sample properties of generalised method of moments estimators. *Econometrica*, *50*, 1029–1054.

Harvey, A., & Phillips, G. (1982). Testing for contemporaneous correlation of disturbances in systems of regression equations. *Bulletin of Economic Research*, *34*(2), 79–81.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica*, *46*, 1251–1271.

Hausman, J., & Taylor, W. (1981). Panel data and unobservable individual effects. *Econometrica*, *49*, 1377–1399.

Helpman, E. (1984). A simple theory of international trade with multinational corporations. *Journal of Political Economy*, *92*(3), 451–471.

Helpman, E. (1987). Imperfect competition and international trade: evidence from fourteen industrial countries. *Journal of the Japanese and International Economies*, *1*(1), 62–81.

Helpman, E. (2006). Trade, FDI and the organization of firms. *Journal of Economic Literature*, *XLIV*, 589–630.

Helpman, E., Melitz, M., & Yeaple, S. (2003). *Export versus FDI* (NBER Working Paper No. 9439).

Henderson, D., & Millimet, D. (2008). Is gravity linear? *Journal of Applied Econometrics*, *23*(2), 137–172.

Intrilligator, M., Bodkin, R., & Hsiao, C. (1996). *Econometric models, techniques and applications* (2nd ed.). New Jersey: Prentice Hall.

Jakab, Z., Kovács, M., & Oszlay, A. (2001). How far has trade integration advanced?: an analysis of the actual and potential trade of three Central and Eastern European countries. *Journal of Comparative Economics*, *29*, 276–292.

Jungmittag, A. (1995). *Langfristige Zusammenhänge und kurzfristige Dynamiken zwischen Direktinvestitionen und Exporten*. Berlin: Duncker & Humblot.

Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, *90*(1), 1–44.

Krishnakumar, J. (1988). *Estimation of simultaneous equation models with error components structure*, *Lecture Notes in Economics and Mathematical Systems*. Berlin: Springer.

Kunze, U., & Schumacher, D. (2003). *Position Ostdeutschlands bei Export nach Polen und Tschechien ausbaufähig* (DIW Wochenbericht Nr. 33/2003).

Lamotte, O. (2002). *The trade potential between Yugoslavia and EU*. Paper presented at the 7th EACES conference ‘Globalisation and Economic Governance’, 6–8 June 2002.

Linders, G., & De Groot, H. (2006). *Estimation of the gravity equation in the presence of zero flows* (Working Paper). Department of Spatial Economics, University of Amsterdam.

Lipsey, R., & Weiss, M. (1981). Foreign production and exports in manufacturing industries. *Review of Economics and Statistics*, *63*(4), 488–494.

Lipsey, R., & Weiss, M. (1984). Foreign production and exports of individual firms. *Review of Economics and Statistics*, *66*(2), 304–308.

Ma, Y., Morikawa, K., & Shone, R. (2000). A macroeconomic model of direct investment in foreign affiliate of Japanes firms. *Japan and the World Economy*, *12*, 953–973.

MacKinnon, J. (2002). Bootstrap inference in econometrics. *Canadian Journal of Economics*, *35*(4), 616–645.

Maddala, G., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, *61*, 631–652.

Markusen, J. (1995). The boundaries of multinational enterprises and the theory of international trade. *Journal of Economic Perspectives*, *9*(2), 169–189.

Markusen, J., & Maskus, K. (1999). *Multinational firms: reconciling theory and evidence* (NBER Working Paper No. 7163).

Matyas, L. (1997). Proper econometric specification of the gravity model. *The world economy*, *20*(3), 363–368.

Mitze, T. (2009). *Endogeneity in panel data models with time-varying and time-fixed regressors: to IV or not-IV?* (Ruhr Economic Papers No. 83).

Nakamura, J., & Oyama, T. (1998). *The determinants of foreign direct investment from Japan and the United States to East Asian countries and the linkages between FDI and trade* (Working Paper No. 98-11). Bank of Japan.

Pain, N., & Wakelin, K. (1998). Export performance and the role of foreign direct investment. *The Manchester School*, *66*(S), 62–88.

Pantulu, J., & Poon, J. (2003). Foreign direct investment and international trade: evidence from the US and Japan. *Journal of Economic Geography*, *3*(3), 241–259.

Park, S. (2005). *Estimating a linear simultaneous equation model with panel data* (Working Paper). Carleton University.

Pesaran, H. (2003). *A simple panel unit root test in the presence of cross section dependence* (Cambridge Working Papers in Economics 0346). Faculty of Economics (DAE), University of Cambridge.

Pesaran, H. (2007). A simple panel unit root test in the presence of cross section dependence. *Journal of Applied Econometrics*, *22*, 265–312.

Pesaran, H., & Yamagata, T. (2008). Testing for slope homogeneity in large panels. *Journal of Econometrics*, *142*(1), 50–93.

Pfaffermayr, M. (1996). Foreign outward direct investment and exports in Austrian manufacturing: substitutes or complements? *Weltwirtschaftliches Archiv*, *132*(3), 501–552.

Plümper, T., & Tröger, V. (2007). Efficient estimation of time invariant and rarely changing variables in panel data analysis with unit effects. *Political Analysis*, *15*, 124–139.

Pöyhönen, P. (1963). A tentative model for the volume of trade between countries. *Weltwirtschaftliches Archiv*, *90*, 93–99.

Prucha, I. (1984). On the asymptotic efficiency of feasible Aitken estimator for seemingly unrelated regression models with error components. *Econometrica*, *52*, 203–207.

Santos Silva, J., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641–658.

Sarafidis, V., & Robertson, D. (2009). On the impact of error cross-sectional dependence in short dynamic panel estimation. *Econometrics Journal*, *12*(1), 62–81.

Sargan, J. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, *26*, 393–415.

Schreiber, S. (2007). *The Hausman test statistic can be negative even asymptotically*. Unpublished manuscript. Goethe University Frankfurt.

Schumacher, D., & Trübswetter, P. (2000). *Volume and comparative advantage in East West trade* (DIW Discussion Paper No. 223).

Serlenga, L., & Shin, Y. (2007). Gravity models of intra-EU trade: application of the CCEP-HT estimation in heterogeneous panels with unobserved common time-specific factors. *Journal of Applied Econometrics*, *22*(2), 361–381.

Song, S., & Jung, B. (2001). Tests for panel regression model with unbalanced data. *Journal of the Korean Statistical Society*, *30*(3), 511–527.

Staiger, D., & Stock, J. (1997). Instrumental variables with weak instruments. *Econometrica*, *65*, 557–586.

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. Andrews & J. Stock (Eds.), *Identification and inference for econometric models: essays in honor of Thomas Rothenberg* (pp. 80–108). Cambridge: Cambridge University Press.

Tadesse, B., & Ryan, M. (2004). Host market characteristics, FDI and the FDI-trade relationship. *The Journal of International Trade & Economic Development*, *13*(2), 199–229.

Tinbergen, J. (1962). *Shaping the world economy*. New York: The 20th Century Fund.

Türkcan, K. (2007). Outward foreign direct investment and intermediate goods exports: evidence from the USA. *Economie Internationale*, *4Q*, 51–71.

UNCTAD (2008). *Foreign direct investment database*. United Nations Conference on Trade and Development, available at: www.unctad.org.

VGRdL (2008). *Volkswirtschaftliche Gesamtrechnungen der Bundesländer (Regional Accounts for German States)*, available at: https://vgrdl.de.

Wang, Z., & Winters, L. (1991). *The trading potential of Eastern Europe* (CEPR Discussion Paper No. 610).

Zarotiadis, G., & Mylonidis, N. (2005). *FDI and trade in the UK: substitutes or complements?* (Working Paper). Department of Economics, University of Ioannina.

Zwinkels, R., & Beugelsdijk, S. (2010). Gravity equations: workhorse or Trojan horse in explaining trade and FDI patterns across time and space? *International Business Review*, *19*(1), 102–115.